University of Alberta

Who Gets Missed: Coverage in the 1992 Alberta Newborn Screening Program

by

Donald William Spady ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Medical Sciences--Public Health Sciences

Department of Public Health Sciences

Edmonton, Alberta

Spring 1997

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced with the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-21208-4

Canada

University of Alberta

Library Release Form

Name of Author: Donald William Spady

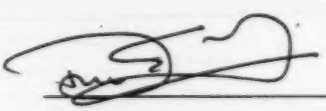
Title of Thesis: Who Gets Missed: Coverage in the 1992 Alberta Newborn Screening Program

Degree: Master of Science

Year this Degree Granted: 1997

Permission is hereby granted to the University of Alberta library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication or other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



Donald W. Spady
176 Quesnell Crescent
Edmonton, Alberta
Canada
T5R 5P3

Date: April 15/97

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Who Gets Missed: Coverage in the 1992 Alberta Newborn Screening Program** submitted by **Donald William Spady** in partial fulfillment of the requirements for the degree of **Master of Science in Medical Sciences--Public Health Sciences.**

D Saunders

Dr. Duncan Saunders (supervisor)

Kyung Bay

Dr. Kyung Bay

F. J. Bamforth

Dr. Fiona Bamforth

Date: APRIL 15/97

DEDICATION

To my family

ABSTRACT

The objectives of this research were to: i) determine which infants, born alive in Alberta in 1992, were not screened within seven days of birth for the diseases congenital hypothyroidism, phenylketonuria, tyrosinemia, and biotinidase deficiency; ii) define the demographic and biological characteristics of infants not screened; and, iii) determine the degree to which infants who had abnormal results were followed up appropriately and expeditiously and a final disposition made. Logistic regression was used to compare demographic characteristics of unscreened infants with screened infants having a high confidence of being accurately matched.

Coverage was calculated to be 98%. Determinants for being unscreened were: death in week one of life, birth weight less than 1500 grams, or between 1500 and 2500 grams, [Reference: over 2500 grams], the single or formerly married maternal marital status [Reference: married], or birth out of hospital [Reference: birth in hospital].

Required repeat screen records were found for only 48 percent of infants. All children with a high phenylalanine level had repeat samples but, repeats for other reasons ranged from 24 to 65 percent.

ACKNOWLEDGMENT

Special thanks go to Jeremy Winter, whose concern about the fate of screened children inspired me to do this project, to my chairman, Dr. Peter Olley, who encouraged me throughout the long process and gave me sufficient time to do it, and to my committee members: Dr. Duncan Saunders, Dr. Kyung Bay, and Dr. Fiona Bamforth for their patience, advice, and encouragement.

I thank Kathy Mazurek in the Department of Clinical Laboratories for helping me to obtain the data, and Dr. Joe D'Almeida and Harry Calhoun of the EPICORE project who helped surmount the problem of transferring the data from laboratory tapes to a personal computer. Very special thanks go to my son, David Spady, without whose programming skills this project would not have succeeded. Finally, to Barbara, who has given me her encouragement, support, and wisdom throughout the project.

TABLE OF CONTENTS

	PAGE
CHAPTER 1	1
Section 1: Introduction and Statement of the Problem	1
Section 2: Literature Review	2
Congenital Hypothyroidism	3
Phenylketonuria	3
Tyrosinemia	4
Biotinidase Deficiency	4
How Good Are the Tests?	4
When Should a Child be Tested?	5
Are There Guidelines for Newborn Screening Programmes?	6
Is Screening Mandatory?	7
Is Screening Cost Effective?	7
How Good is the Coverage?	7
Who Does Not Get Screened?	9
Why Are Some Affected Children Missed?	10
Summary	11
Section 3: The Alberta Newborn Screening Program	12
What Normally Happens	12
Screening for Congenital Hypothyroidism	15
Phenylketonuria and Tyrosinemia	15
Biotinidase Deficiency	16
Repeats Done Elsewhere	16
Coverage	16
Section 4: Project Aim and Objectives	18
CHAPTER 2: METHODS	19
Section 1: Overview	19
Section 2: Data Preparation	20
Section 3: The Match	24

Section 4: Statistical Analysis 27
Completeness of Coverage 27
Program Performance in Infants Who Were Screened 28
Analysis of "Irregular Screening Events" 28
CHAPTER 3: RESULTS 29
Section 1: The Data 29
Vital Statistics Data 29
The Screen Data 32
Section 2: Coverage 33
Variation Among Health Regions 34
Infants Born Out of Hospital 34
Characteristics of the Good Match, Uncertain Match, and Not Matched Groups 34
Birthweight 36
Maternal Marital Status 36
Maternal Age 41
Relationships Between Key Variables and Match Status 41
Logistic Regression Analysis of Matched vs Not Matched Infants 43
Section 3: Program Performance 45
Blood Sources 45
Timing of Initial Blood Samples 45
Blood Samples Requiring Repeat 48
The Non-Repeaters 52
The Repeaters 54
Section 4: Specific Diseases 55
Phenylketonuria 55
Biotinidase Deficiency 55
Tyrosinemia 56
Congenital Hypothyroidism 56

NSQ/US Samples 57
Samples Obtained Before 24 hours of Age 57
Repeat Samples When the Original was Normal 57
CHAPTER 4: DISCUSSION 58
Section 1: Data Quality 58
The Source of the Data 58
The Accuracy of the Match 59
The Creation of Specific Population Subgroups 60
The Detection of Repeat Samples 61
The Assignment of Age at Blood Sampling 61
The Value of this Research 62
Section 2: The Results 64
Coverage 64
Who Got Missed 66
Timing of Initial Samples 69
Repeat Screens 70
Section 3: Specific Results 75
Phenylketonuria 75
Congenital Hypothyroidism 75
NSQ/US Results 76
Tyrosine 76
Biotinidase Deficiency 77
Repeat Samples That Were Not Indicated 77
Section 4: Conclusions 79
Some Suggestions Regarding the Alberta Newborn Screening Program 80
REFERENCES 81
APPENDIX 1: DATA ABSTRACTION 85
APPENDIX 2: THE MATCHING PROCEDURE 92

LIST OF TABLES

TABLE		PAGE
1-1	Typical Messages Found on Abnormal Screening Laboratory Results	14
2-1	Description of the Variables in the Screen Data File	21
2-2	Description of the Variables in the Birth File	22
2-3	Available Variables for Linking and for Assessing Missing Infants	23
2-4	Summary of Matching Criteria, Confidence Levels, and Match Progress	26
3-1	Descriptors of the 1992 Birth Cohort	30\
3-2	Total Births, Stratified by Health Region of Birth	31
3-3	Source and Number of Screen Records	32
3-4	A Numerical Description of Key Variables	35
3-5	Characteristics of Infants Born Out of Hospital	36
3-6	Demographic Characteristics Stratified by Match: Continuous Data	37
3-7	Demographic Characteristics of the NM, GM, and UM Groups: Place of Birth and Mobility	38
3-8a	Frequencies of Birth weight in All Children	39
3-8b	Frequencies of Birth weight in Children Alive at Age 1 or Dying After Age 7 Days	39
3-9	Match Status by Marital Status of the Mother	40
3-10	Mother's Age by Age Category and Match Status	41
3-11	Marital Status, Birth weight, and Age Relationships	42
3-12	Logistic Regression: Not Matched Compared to Good Match	44
3-13	Characteristics of Infants with Blood Taken When Over 7 Days of Age	45
3-14	Comparison of Infants by Timing of Blood	47

3-15 Reasons for the Need to Repeat a Screen Sample 48
3-16 Characteristics of Good Match Infants Who Require Repeat Screen 49
3-17 Characteristics of Infants Needing Repeat Screen: Categorical Data 50
3-18 Logistic Regression: Infants Needing Repeat vs Those Not Needing Repeat. Good Match Infants Only 51
3-19 Good Match Infants Who Had Repeat Screens Compared to Those Who Did Not Have Repeat Screens 52
3-20 Good Match Infants With Repeat Screens Compared to Infants Without Repeat Screens: Categorical Data 53
3-21 Logistic Regression: Infants Not Having Required Repeat vs Those Having Required Repeat 54

LIST OF FIGURES

FIGURE		PAGE
1-1	What Happens to a Newborn Screen	13
2-1	The Match Procedure	24
2-2	Data Combinations for Analysis	27
3-1	A Summary of the Results	33

LIST OF ABBREVIATIONS

ABBREVIATION	MEANING
AHCIP	Alberta Health Care Insurance Plan
BD	Biotinidase deficiency
CBD	Correct birth date
CH	Congenital Hypothyroidism
DE	Decreased
GM	Good match
NM	Not matched
NSP	Newborn Screening Program
NSQ/US	Not sufficient quantity/Unsatisfactory sample
OOH	Out of hospital
OR	Odds ratio
PKU	Phenylketonuria
SCRID	Screen ID Number
SPSS	Statistical Package for the Social Sciences
TSH	Thyroid stimulating hormone
T4	Serum thyroxine
UM	Uncertain match
VID	Vital Statistics ID Number
WMC	Walter Mackenzie Centre

CHAPTER 1

Section 1: Introduction and Statement of the Problem

Since the introduction of the Guthrie test for phenylketonuria (Guthrie & SCSI, 1963) in the early 1960's, the screening at between three and seven days of age of all newborn infants for certain congenital metabolic diseases has become standard practice in most of the economically developed world. Each year more than seven million infants are screened for metabolic disease and over 1200 children avoid the tragedy of preventable subnormal intelligence or profound mental retardation (Willi & Moshang, 1991).

A common concern in all newborn screening programs is the completeness of coverage; i.e. what proportion of children who should have been screened were screened. This is commonly estimated by comparing the number of unique screen samples analysed with the number of births registered in the same time period and the same geographic region. However, rarely is a formal link made between the two, and even more rarely is any attempt made to determine which children are being missed.

In North America, it has recently become common for children to be discharged shortly after birth, often when the child is only 24 hours old. This practice has increased the possibility of a child not being screened before discharge as well as the probability of a child having an incorrect result because a blood sample was taken too soon. As a consequence, some affected children may not be detected and treated promptly and other children may require a second blood sample because of false positive results.

These issues of coverage and the potential decrease in coverage due to early discharge point to the need to explore in detail how completely newborn screening programs cover the eligible population and to determine if there are groups of infants that are at particular risk for non-screening. This project will address some of these issues. It will assess the coverage of the Alberta Newborn Screening Program for the year 1992, a time when early discharge was relatively uncommon, and describe the characteristics of children who were never screened. As well, it will assess the effectiveness of follow up of children who have initial screen results that require verification.

Section 2: Literature Review

Last (1995) defined screening as:

> *"The presumptive identification of unrecognized disease or defect by the application of tests, examinations or other procedures which can be applied rapidly. Screening tests sort out apparently well persons who probably have the disease from those who probably do not. A screening test is not designed to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and necessary treatment."*

Key features of an effective screening program include: (i) using accurate, reproducible, and sensitive diagnostic tests, (ii) screening all eligible subjects; (iii) ensuring rapid follow up and verification of 'positive' results; (iv) the likelihood of a sufficient yield of otherwise undetected patients; (v) promptly providing and instituting definitive therapy; and (vi) appropriate follow up of suspect individuals (Wilson & Jungner, 1968).

Newborn screening programs are among the most effective screening programs found in health care today. In such programs, all newborn infants are screened for specific diseases at between one and seven days of age. The disorders most commonly screened for are congenital hypothyroidism and phenylketonuria. Other metabolic disorders, such as sickle cell disease, biotinidase deficiency, tyrosinemia, congenital adrenal hyperplasia, galactosaemia, and maple syrup urine disease, are also included in some screening programs (Newborn Screening Committee, 1995). These diseases share the features of being rare, present at birth, very hard or impossible to detect clinically in the newborn period, capable of profoundly and permanently impairing the health and mental development of the child from shortly after birth if not treated, and amenable to appropriate therapy, preferably starting within ten days of life and resulting in normal health and development. In a few jurisdictions, infants are also screened for hearing disorders, but this has not as yet gained widespread acceptance. This project is particularly concerned with the four disorders screened for by the Alberta Newborn Screening Program [NSP]: phenylketonuria (PKU), congenital hypothyroidism (CH), biotinidase deficiency (BD), and tyrosinemia (TY).

Congenital Hypothyroidism

Congenital hypothyroidism, a disorder of thyroid metabolism, affecting about 1:4000 infants, is one of the most common causes of preventable mental impairment in children. Several variants of CH exist but key to all of them is the early detection of the presence of disease and the institution of therapy as soon as possible after birth. With early therapy most children with CH will have normal intellectual development and the effects of the more severe forms of CH will be minimized. The physical findings associated with CH are non-specific, subtle and easily missed and early diagnosis is normally dependent on laboratory measurement of thyroid function. Screening for CH was introduced in 1974 (Dussault, Coulombe, Laberge, et al, 1975). This disease is detected by measurement of thyroid stimulating hormone (TSH), or serum thyroxine (T4) in blood. T4 is the most common primary measure used in the United States with TSH as a secondary measure (LaFranchi, Dussault, Fisher et al, 1993). Many Canadian, European, and Japanese programmes use TSH as the primary measure (Gruters, Delange, Giovannelli et al, 1994). Newer TSH assays are more sensitive than before and may become the primary measure in many US states; however, TSH values are higher in the first 24-48 h than later for all infants, and with the increased frequency of early discharge from hospital, there may be more false positive results.

Phenylketonuria

Phenylketonuria is a disorder of amino acid metabolism that affects about 1:15000 children and results in moderate to severe mental retardation if diagnosis and treatment are not instituted as soon as possible after birth. It was the development of the Guthrie test in the early 1960's to detect PKU in newborn infants (Guthrie & SCSI, 1963) that stimulated the development of mass newborn screening programs. Many agencies still use the Guthrie test to detect high levels of phenylalanine, but other methods exist, and with the development of thin layer chromatography, this became a common technique. As with congenital hypothyroidism, the earlier definitive therapy is instituted, the less the likelihood of the child being mentally retarded.

Tyrosinemia

The finding of a high tyrosine level is a relatively common finding among newborn infants and often has little clinical significance; however, high tyrosine levels are also found in tyrosinemia, of which there are several variants. Tyrosinemia Type I is associated with severe liver disease and treatment consists of diet therapy as a stop-gap and liver transplant when possible. Screening for this disease alone would be of little benefit. Tyrosinemia type II, while rare (1:50000), is associated with mild-moderate forms of mental retardation and is amenable to diet therapy. A third disorder, transient tyrosinemia, characterized by transiently high levels of tyrosine, may be associated with mild mental retardation (Scriver, Beaudet, Sly, Valle, 1989; Mamunes, Prince, Thornton, et al, 1976) and can be treated with high doses of Vitamin C. This seems to be most common among premature infants. Screening for tyrosinemia is usually a "spinoff" from screening for PKU using thin layer chromatography, which analyses several amino acids simultaneously.

Biotinidase Deficiency

Biotinidase deficiency was first described by Wolf et al (Wolf, Grier, Allen, Goodman, Cein, 1983). It is an inherited metabolic disorder that can be asymptomatic, or can produce skin rash, alopecia, seizures, ataxia, developmental delay and death. Early detection and subsequent treatment with biotin may resolve the problem (Widhalm, Wintersperger, Bischof, Brix, 1995). Screening for this rare (1:60000) disease is done by measurement of biotinidase on a filter-paper blood specimen. Timing is not as critical as with CH, PKU or TY; however, the disease does present in infancy.

How Good Are the Tests?

The sensitivity and specificity of the tests vary, but are generally very high. For congenital hypothyroidism, the sensitivity of the test for TSH has been reported to be 97 percent and the specificity is 99.9 percent (Pharoah & Madden, 1992), but these figures depend upon the level of TSH considered as abnormal. For PKU the respective figures are 100 percent and 99.9 percent (Hisashige, 1994). This means that in most cases, if a child is affected with these diseases, the tests will detect them and conversely, if the test is

negative, the child is very unlikely to be affected. In some instances, however, biologic variants of the disease in question may result in falsely negative results, and thus a missed diagnosis (MacMillan & Mabry, 1995). The tests also vary in sensitivity and specificity with respect to the age of the infant. Testing before 24 hours of age is associated with increased numbers of false positive results (Foley, Fisher, Shapiro et al, 1987; Saslow, Post, Southard, 1996) for CH and possibly increased numbers of false negative results for PKU (Holtzman, McCabe, Cunningham, et al, 1981; Sinai, Kim, Casey, Pinto-Martin, 1995). Premature infants are also more likely to have a false positive CH result than are full term infants; thus they should be tested twice, once in the first week of life and again before discharge (Gruters, Delange, Giovanelli, et al, 1994).

When Should a Child Be Tested?

Screening should take place between one and seven days of age. When first instituted, screening samples were obtained at about three to five days. This time was chosen because the Guthrie test for PKU depended on the child having been fed phenylalanine, an amino acid found in milk. Infants who had not been fed may well have a false negative result. Current methods of measurement for PKU are more sensitive and thus the requirement for a milk feeding before sampling is less important; however, testing before 24 hours of age may still provide false negative results, and it is still best to wait until the child is at least 24 hours old before sampling. Screening for congenital hypothyroidism should occur after 24 hours of age as the values of T4 or TSH may be artificially low (T4) or high (TSH) before this time (Foley, Fisher, et al, 1987; Dobbin, 1987). If screening is delayed beyond seven days, the diagnosis, and essential treatment, may be unnecessarily delayed. For children affected with CH or PKU, it is considered desirable that they receive definitive therapy by 21 days of age or earlier.

If a positive result ensures, the parent must be informed and a repeat sample taken. All positive results require verification by re-screening and the more "false positives" there are, the greater the number of infants needing unnecessary follow up and re-screening. As well, the anxiety of a parent waiting for that second result is great (Rothenberg & Sills, 1968; Fyro & Bodegard, 1987; Sorenson, Levy, Mangione, et al, 1984;).

Some hospitals may discharge infants before 24 hours of age and before being screened. The parent may be given a screening form and told to have the screening done sometime in the next few days. While this practice is becoming less common, it raises the possibility that the infant will not be screened at all. Some hospitals screen twice, the first time at less than 24 hours and again at three to five days when the child is at home. This increases the cost of screening and also introduces the need to match two samples. As well, with samples taken in the hospital environment, usually one or two laboratory personnel take all of the samples. In the "home" environment, however, many different technicians may take samples; this can increase the likelihood of poor samples being obtained.

Are there Guidelines for Newborn Screening Programmes?

To ensure proper performance of newborn screening, guidelines have been developed describing the institution and management of a newborn screening program. One of the most complete sets of guidelines is that produced by the Committee on Genetics (1992) of the American Academy of Pediatrics, but other agencies around the world have also prepared guidelines (Gruters, Delange, Giovannelli et al,1994; LaFranchi, Dussault, Fisher et al, 1993). While these guidelines differ in specifics and emphasis, they all emphasize the concepts of: (I) centralization; (ii) rigid coverage of all infants at an age of greater than 24 hours and less than seven days; (iii) the use of accepted screening tests for the respective metabolic diseases; (iv) a clear line of responsibility for transmission of all screening results to the physician or hospital identified on the filter paper, and a more direct and immediate method of informing physicians of abnormal results; (v) follow up of suspect results until such time as a diagnosis is made and definitive therapy is instituted; (vi) some form of ongoing evaluation, usually participation in national or international quality control programs that allow an assessment of sample processing; and, (vii) keeping proper records. Some agencies also emphasize education of the population with respect to newborn screening.

Is Screening Mandatory?

Some jurisdictions, particularly states in the United States, have legislation mandating a newborn screening practice. In other jurisdictions, such as Alberta and most of the Canadian provinces, screening programs are voluntary. It is unclear as to which is better, as both types of program appear equally effective in assuring coverage. Regardless of whether the program is legislated or not, the institution of a newborn screening program in any jurisdiction constitutes an implied mandate and pledge that every eligible infant liveborn will be tested and, if 'positive', will receive appropriate diagnostic testing and therapy.

Is Screening Cost Effective?

Estimates of the cost benefit of newborn screening vary but are in the range of 1:2.5 to 1:6.6 for PKU and up to 1:13.8 for CH (Hisashige, 1994; Dhondt, Farriaux et al, 1991). In other words, for every dollar spent in a NSP, somewhere between $2.5 and $13.8 are saved by finding and treating the affected child. The higher ratio for CH is because the costs of long term therapy are significantly less for a child with CH than they are for a child with PKU. Hisashige (1994) estimates the net benefit to detect one case of PKU is about Can$390,000. For diseases with a higher incidence than PKU, the net benefit would be even greater.

How Good is the Coverage?

Coverage refers to the proportion of children actually screened relative to all children born alive in the same time period. The estimates of coverage are based on comparing the number of screening tests performed in a given year with the number of children born in the same period, after allowing for known repeat blood samples and samples from children born out of province. Estimates of completeness vary from country to country and year to year, but generally, once start-up and administrative problems are worked out, estimates of coverage often approach 100 percent. Thus, for example, in Germany, 94 to 95 percent of newborns are screened (Schoenberg & Klett, 1987), in Australia about 99 percent are screened (Connelly, 1987), in Japan 99.5 percent (Irie, Nakajima, Inomata, et al, 1987), and in Italy, 47 to 99 percent, depending on

the region serviced (Italian Committee, 1987). In the United States, coverage for 1992 varied from about 76 percent for Nevada to 100 percent for several other states; however, Puerto Rico, an American Trust Territory, screened only eight percent of its newborn population (Newborn Screening Committee, 1995).

Coverage has been assessed in ways other than simple comparisons of birth and screening numbers. On a few occasions, a direct link to registry data has been tried. Griffiths, Morris et al (1987) describe a NSP in Birmingham UK, which used electronic communication between regional hospitals and a central laboratory to ensure that every birth was recorded. When a screening blood sample was received, it was linked to the appropriate birth record. With this method, the agency knew which infants were not screened, and often could determine the reason for non-screening, such as death, moved, or refused. Smith, Cook, and Beasley (1991) refer to a MSc thesis by Hunter that compared matched birth registry data to screening data in an inner London borough and found coverage to be only 93.5 percent.

Pharoah and Madden (1992) reviewed the newborn screening program for CH for the years 1983 to 1989 in the Mersey region of the UK. Birth registration data that had been entered into a child health computer file were used as the denominator in the statistics. Screening results were normally entered into the same computer file thus allowing a link to birth record and determination of who was not screened. Coverage ranged from 96.2 percent to 100 percent, but there was failure to follow up infants who had not been screened. Reasons for non-coverage were not well described but appeared to be mainly administrative and included attribution of the screening result to the administrative area of birth, rather than the administrative area of residence, and lack of recognition of a blood sample as a repeat. In three regions, computerization of records had not occurred and no estimate of coverage was made. Sixty infants were found to be hypothyroid; all were receiving definitive therapy by age 26 days. This study demonstrated that a birth-screen link was possible and that it could work effectively if appropriate measures were taken to ensure proper identification of the child, his or her birthplace and normal place of residence.

In a New York study from 1988 data, live-birth records for a random sample of 1000 infants born to New York State residents were matched, either by computer or by hand, to newborn screening records (Pass, Schedlbauer, MacCubbin, Glebatis, 1991). Screening specimens were obtained for 985 infants; eight of the remaining were never obtained or lost in transit and seven reflected infants born out of state and screening status could not be determined. Thus, screening coverage was at best 99.2 percent and at worst 98.5 percent. Black infants from New York city and low-birthweight infants were represented disproportionately among those not tested.

These studies are the only ones the reviewer could find that tried to link birth data to screening data and, in correspondence with directors of screening programs around the world, no other agencies described doing this. The only other study is the one described by this project.

Who Does Not Get Screened?

None of the studies linking birth and screen data described above examined possible risk factors as to why infants were missed. In an attempt to determine which children were at risk for not being screened, Streetly, Grant et al, (1994) compared birth records with screening records for infants born in a 3 month period and resident in a district of London. They found that infants who moved into or out of the region after birth were at greater risk of not being screened than were infants who were born in the region and of parents who normally lived there. Differences in ethnic status appeared to have a bearing on screening coverage, but these disappeared after accounting for mobility and sub-district. Possible reasons for non-screening were mobility around the time of birth, admission to a special care unit, twin birth, obstetric complications, birth out of country, and refusal. The authors noted that no one person had responsibility for ensuring adequate coverage and record maintenance and they recommend that a system for monitoring test results be maintained, much the same way as is the system for monitoring immunization status.

Why Are Some Affected Children Missed?

It is almost inevitable that some children with disease will be missed in screening programs. Sometimes this will be due to the vagaries of the disease or to the fact that the screening tests are not perfect (MacMillan & Mabry, 1995). It is estimated that in screened populations, between 1:50 and 1:80 cases of PKU and 1:120 cases of CHT are missed (Holtzman, Stazyk, et al, 1986; Verkerk, PH, Vaandrager GS, Sengers RC, 1990), either because the infant was never screened, the initial screen was negative, there was laboratory or clerical error, follow up was incomplete, or the disease was a hard to detect variant. This estimate of 1:120 is much less than that of Fisher (1987) who estimated that as many as 1:10 cases of CH may be missed. He estimates that:

> *"...3 infants with hypothalamic-pituitary TSH deficiency, perhaps five (range 3-14) infants with a delayed rise in serum TSH, and perhaps four (range 3-6) infant victims of errors in sampling or sample processing will be missed for every 100 infants detected with congenital hypothyroidism; this would translate to ... 1 infant missed per 37500 infants screened. ".*

In all likelihood, the most important determinant of an affected child not being detected in infancy is that the child was never screened, the blood sample obtained was misinterpreted, was unsatisfactory, or it was obtained too early in the child's life, or follow up for repeat blood samples was incomplete (Fisher, 1987; Leger, 1990; Sinai et al, 1995). Sometimes it is lack of knowledge of the appropriate guidelines. For example, Sinai (1995) described a situation where 24 percent of newborns were discharged by 24 hours of age. Many of these infants did not have a repeat screen for PKU, apparently because pediatricians and other physicians were unaware of the guidelines for newborn screening and thus for the need to obtain a repeat. Apart from the personal tragedy that these missed cases represent, they also constitute a significant legal concern. In the United States, up to 1986, of 76 'missed' cases, 29 percent resulted in legal action and settlements ranged from 1 million to 20 million dollars (Holtzman, Slazyk, Cordero, Hannon, 1986).

As well, plain, simple, errors occur and all of them interfere with the effectiveness of the screening program. Some errors are inherent in any newborn screening program.

These include: not sampling every child, obtaining a sample before 24 hours or after seven days of age, obtaining an unsatisfactory blood sample, incorrectly labelling the sample, incorrectly transporting and storing the sample before analysis, keeping the sample longer than necessary at a peripheral site, incorrect data entry of the sample at the analytical laboratory, incorrect analysis of the sample or interpretation of sample results, not notifying the responsible individual of abnormal results, and failure to ensure that abnormal results or unsatisfactory samples are repeated within an appropriate time frame. Some errors are peculiar to specific hospital practices, notably early discharge from hospital. These errors relate mainly to the problems associated with interpretation of results of samples obtained before 24 hours of age and the requirement to follow and retest ALL children who had blood samples taken before 24 hours of age.

Summary

In summary, newborn screening programs are demonstrably effective mechanisms to detect specific metabolic disorders which are present at birth and amenable to therapy. The methods used to detect disease are highly sensitive and specific. Nevertheless, children with detectable disease still get missed. Although some of the missed diagnoses may be due to variations in the diseases in question, the principal reasons for children not being detected appear to be incomplete coverage of the entire population of newborn infants, laboratory error, and incomplete follow up of suspect results. In recent years the possibility of a child being missed has become more likely because of the discharge of infants early after birth, at a time when the detection of disease is not optimal. While screening programs may closely monitor their laboratory procedures, most programs do not audit their procedures to determine how complete their coverage is or to define the demographic and social characteristics of which children are never screened.

Section 3: The Alberta Newborn Screening Program

In Alberta, a newborn screening program has been active since 1967 when screening for phenylketonuria, a disorder of amino acid metabolism, was started. The screening program has since expanded to cover congenital hypothyroidism, biotinidase deficiency, and tyrosinemia. The Alberta Newborn Screening Program (NSP) is administered by the Alberta Hereditary Diseases Program and is carried out in the Department of Laboratory Medicine of the Walter Mackenzie Centre (WMC) of the University of Alberta under the direction of Dr. Fiona Bamforth. All initial newborn screening samples taken in the province are sent here for analysis and each year approximately 42000 infants are screened for PKU, CH, BD and TY. As in most provinces of Canada, the screening of a newborn infant is not mandated by law. Therefore, the success of the program depends upon the cooperation between the hospital, health care professionals, and the parent.

What Normally Happens?

In practice, in 1992 most infants born in Alberta were screened at two to three days of age. Some infants, notably those born at the Royal Alexandra Hospital and Misericordia Hospitals in Edmonton, were part of a program whereby they were discharged at age 24 hours. While many of these infants may have had a sample taken before 24 hours of age, they were all supposed to be followed up by public health nurses and a repeat sample obtained before seven days of age. For hospitals with a large number of newborns, samples were to be sent daily to WMC for analysis. For hospitals with fewer newborns, samples could be sent every other day or even less commonly, but in all instances, the samples were not to be retained for more than three to four days before being forwarded to WMC. Figure 1-1 is a schematic flowsheet illustrating the general process employed by the Alberta NSP in 1992. Upon receipt by the NSP, the samples were assigned a laboratory number, pertinent demographic information was entered into the laboratory computer, and the sample prepared for analysis. All samples were analysed and all analyses had a report returned to the referring physician. Even those samples considered unsatisfactory (NSQ/US) were analysed, if it was at all possible. If

the result of an NSQ/US sample was "normal", the samples were reported as NSQ or US and a request was made for a repeat sample to be provided. If the result of an NSQ/US sample was "abnormal", it was reported as such to the physician, together with a request for a repeat. This practice of analysis of these NSQ/US samples has resulted in at least one child being detected with congenital hypothyroidism.

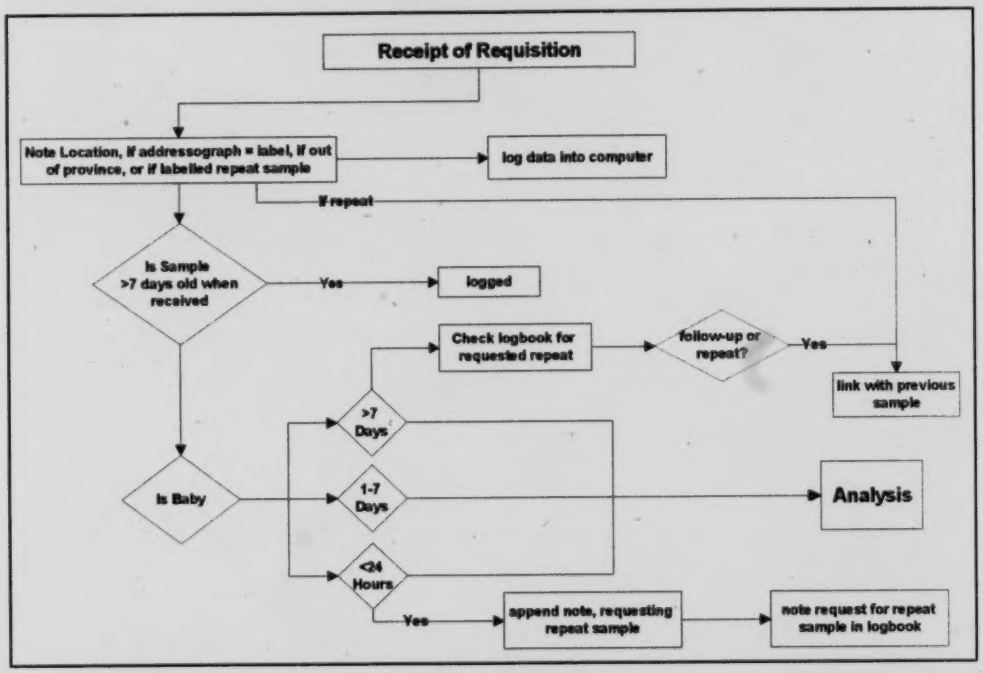


Figure 1-1 What Happens to a Newborn Screen

The program informed the physician of all normal results by sending the physician a copy of the screening report. An abnormal result was usually conveyed to the physician by telephone, with the request that a repeat sample be obtained. Table 1- 1 shows some of the messages found on abnormal screening laboratory reports. Less significant abnormal results were usually sent by mail, with the report stating that there was an abnormal finding and that the child should have a repeat sample in about a month. This was most likely to occur when high levels of tyrosine are found. It could have occurred with thyroid screens that showed a moderately high TSH but a normal T4, reflecting a

Table 1-1. Typical Messages found on Abnormal Screening Laboratory Reports

As stored in Data file	As found in raw data
'SL' (slightly increased amino acid)	SLIGHT TO MODERATE GENERALIZED INCREASE IN BLOOD AMINO ACIDS OFTEN OF NO SIGNIFICANCE BUT COULD INDICATE GENERAL ILLNESS, CATABOLISM, OR LIVER DISEASE. SUGGEST CHECK CLINICALLY AND REPEAT NEONATAL METABOLIC SCREEN.
'TY' (tyrosine)	1. TYROSINE INCREASED. USUALLY A BENIGN, TRANSIENT CONDITION, OFTEN ASSOCIATED WITH PREMATURITY. SHOULD BE FOLLOWED UP TO EXCLUDE LIVER DISEASE OR PERSISTENT HYPERTYROSINEMIA. SUGGEST REPEAT SCREEN IN ABOUT 1 MONTH OR SOONER IF CLINICALLY INDICATED. SUCCINYLACETONE NO LONGER MEASURED ROUTINELY AS PART OF THE NEWBORN SCREEN. ASSAY IS STILL AVAILABLE WHERE THERE IS CLINICAL OR BIOCHEMICAL SUGGESTION OF TYROSINEMIA TYPE I. 2. **(repeat sample)** TYROSINE STILL SLIGHTLY ELEVATED AS ON SAMPLE COLLECTED 00/00/94. SUGGEST FOLLOW UP WITH PLASMA AMINO ACID QUANITITATION AND SUCCINYLACETONE (2 ML HEPARINIZED BLOOD REQUIRED)
'PH' (phenylalanine)	MARKED HYPERPHENYLALANINEMIA. IMMEDIATE CLINICAL AND BIOCHEMICAL FOLLOW UP IS RECOMMENDED.
'HI'	1. TSH SLIGHTLY ELEVATED BUT NOT IN THE RANGE USUALLY SEEN IN PRIMARY HYPOTHYROIDISM. SUGGEST CLINICAL FOLLOW-UP AND SERUM THYROID FUNCTION TESTS. 2. BORDERLINE TSH (15-25 MIU/L). T4 SLIGHTLY DECREASED. SUGGEST REPEAT NEONATAL SCREEN TO EXCLUDE PRIMARY HYPOTHYROIDISM. 3. INCREASED TSH STRONGLY SUGGESTIVE OF HYPOTHYROIDISM ALTHOUGH T4 IS NORMAL. SUGGEST CLINICAL ASSESSMENT AND FOLLOW-UP WITH SERUM THYROID FUNCTION TESTS.
'DE' (Biotin)	1. BIOTINIDASE ACTIVITY DECREASED. PLEASE SEND REPEAT NEONATAL SCREEN. 2. REDUCED LEVEL OF BIOTINIDASE ACTIVITY. UNLIKELY TO BE OF SIGNIFICANCE. SUGGEST REPEAT SCREEN IN ONE MONTH.
'NSQ/US'	1. BLOOD CLOTTED AND LAYERED ON FILTER PAPER. WE WERE UNABLE TO ELUTE BLOOD FOR NEONATAL SCREEN. PLEASE SEND A REPEAT SAMPLE. 2. PLEASE SEND A REPEAT SAMPLE. PLEASE FILL AND SATURATE THE CIRCLES ON THE REQUISITION.
Less than 24h	INFORMATION ON REQUISITION INDICATES THAT SAMPLE WAS COLLECTED BEFORE 24 HOURS OF AGE. THIS IS TOO EARLY FOR RELIABLE DETECTION OF PKU. SUGGEST REPEAT SCREEN.

child with likely normal thyroid functioning but one which should be verified.

The names of all infants who had screen results indicating the need for a repeat sample were entered into a log and the name was checked off when a repeat sample was obtained. The laboratory identified repeat samples if they were so labelled and also by noting the age of the child when the blood was taken. If the age was over seven days, the name on that sample was compared to the names on the "repeat log" to see there was a

match. If no match was found, then the sample was considered a new sample. If a match was found, then it was linked to the appropriate name. If, by chance, a repeat sample was obtained before the infant was seven days old, then it may not have been recognized by the lab as a repeat. The "repeat log" was, and is, a static document in that there is no time frame after which a reminder is sent to the physician to obtain a repeat sample.

The NSP keeps a record of infants diagnosed with any of the diseases screened for but does not maintain a long term follow record of these children or their progress.

Screening for Congenital Hypothyroidism

In 1992, TSH was the primary measure used to detect CH. When a child had an abnormal TSH, a T4 was determined, usually using the same blood sample. Values of TSH of >25 mU/L were considered suspect and required follow-up to verify the presence or absence of disease. If the T4 was within a normal range, the possibility of CH was decreased but not eliminated and a repeat TSH and T4 should have been obtained before the child was one month old. At the same time, a request was sent to the physician for a repeat sample within one month. If the result was highly suggestive of CH (TSH >50), the physician was contacted by telephone and advised to obtain repeat serum samples immediately. These repeat samples may have been done in a private laboratory and the screening program did not know for certain if they were done or what the results were.

Phenylketonuria and Tyrosinemia

Phenylalanine was measured using manual thin layer chromatography and the result reported as "normal" or a message was provided stating the type of abnormal result obtained. This could have been a high phenylalanine level, a high tyrosine level, or some other, less common amino acid abnormality. In every instance, an abnormal result was followed by a request to obtain a repeat blood sample.

In Alberta, tyrosine was screened for, presumably because of the potential benefit to the child, but also because it is a byproduct of the amino acid screen for PKU. High levels, when found, were reported to the physician and a repeat sample requested. For 1992 only, there was a pilot project where, if a high level of tyrosine was found, a measure of succinylacetone was done on the same sample. This compound, when present

in raised amounts, suggested the diagnosis of Tyrosinemia Type I. This project was stopped at the end of 1992.

Biotinidase Deficiency

Biotinidase was measured using a colorimetric technique. When the screening result showed a reduced level of biotinidase, a repeat sample was requested to be sent within a month. Absent biotinidase activity was followed up by a request for an immediate repeat newborn screen.

Repeats Done Elsewhere

Repeat samples may also have been analysed at the Biochemical Genetics Laboratory of the Alberta Children's Hospital. Where samples were identified as repeats or when an abnormal PKU, or Biotin, or Tyrosine result was obtained by the Calgary centre, a copy of the result was sent to the WMC Laboratory in Edmonton. If the sample was not identified as a repeat and was not otherwise identified as a newborn screen, no report would be sent to Edmonton. Thus, a child with an initially abnormal tyrosine value would be listed in the WMC log as requiring a repeat. The child may have had the repeat done in Calgary. If it was normal and not recognized as a repeat, the WMC would not have been informed and the log of abnormal results not changed to reflect that follow up was completed. If the child's repeat was not done, no one would be aware of the incomplete follow-up.

Coverage

Coverage in Alberta has never been formally audited. Estimates are determined by matching samples received with birth demographics. In Alberta in 1992, when infants were screened at two to three days, it was estimated by the NSP that about 98.3 percent of the 42,392 infants born were screened; this figure was comparable to previous years. Twelve children were detected with congenital hypothyroidism. As well, two children with PKU, six with hyperphenyalaninemia, a less serious problem of phenylalanine metabolism, and one with BD were also detected.

In 1994, when this project started, while there was a standard policy with respect to screening early discharged infants in Alberta, each hospital had its own "usual practice".

Infants may not have been screened before discharge and thus may never have been screened or may have been screened too late. In some cases, the parents were given a Newborn Metabolic Screen requisition and told to take it to their doctor when the infant was two or three days old. Sometimes they would be told that a public health nurse would visit and take the blood. How common these "out of hospital" collections occurred is unknown. As well, there was no mechanism in place to ensure screening for infants born at home, thus they constituted another potentially high risk population.

Section 4: Project Aim and Objectives

The above is an outline of the issues faced when considering newborn screening programs. While these programs appear effective in detecting most infants having specific metabolic disease, there remain the problems of: i) an inconsistent screening policy for infants discharged early; ii) an unknown number of infants possibly not being screened; iii) infants being screened at too early or too late an age; iv) infants being screened and a false positive result obtained; and, v) incomplete follow up of positive results. This project addresses these issues as they relate to the Alberta Newborn Screening Program in 1992.

The aim of this project is to document the usual practice in 1992 of the Alberta Newborn Screening Program. There are three objectives:

1. To determine the proportion of liveborn infants who were not screened within seven days of birth for the diseases phenylketonuria, congenital hypothyroidism, tyrosinemia, and biotinidase deficiency in Alberta during the year 1992.
2. Using available demographic and biological data, to define the characteristics of infants not screened, or who could not be identified with confidence as being screened.
3. To determine the degree to which infants who, when screened and had results requiring follow up, were followed up appropriately and expeditiously and a final disposition made, and to define the characteristics of infants who were not followed up.

The outcomes of interest are: (i) the completeness of coverage of the screening program and the timing of blood sampling (Objective 1); (ii) a description of the available characteristics of screened vs non-screened infants (Objective 2) and (iii) the age at which initial follow up was completed for those for whom data are available, the proportion of follow up requests that reflected a false positive on initial screening, and characteristics of infants who had false positive results and of infants lost to follow-up (Objective 3).

While the project has the objectives outlined above and is mainly descriptive, several hypotheses arise out of the studies mentioned above that can be tested with the available data. The research hypothesis is that the Alberta Newborn Screening Program ensures complete, efficient, and accurate screening of infants for the metabolic diseases phenylketonuria, congenital hypothyroidism, tyrosinemia, and biotinidase deficiency.

CHAPTER 2: METHODS

Section 1: Overview

The method used to meet the objectives was to link birth record data with existing screening record data, to determine those infants who had never been screened and to compare existing biologic and demographic data of these infants with those who were screened. The specific objectives have been defined earlier. For screened infants, the age when blood was obtained was calculated and the results of the screens were tabulated. For screened infants who required a repeat screen, the completeness and timeliness of rescreening was determined, as were demographic and biologic factors that described groups of rescreened with non rescreened infants.

Objective 1 was addressed by matching all births happening in Alberta in 1992 and registered with Alberta Registries (Vital Statistics), the "Gold Standard" of who was eligible for screening, with the equivalent Newborn Screening Program (NSP) data to describe the percentage of infants born in 1992 who were screened. Infants unmatched and otherwise not accounted for were considered to be unscreened.

Objective 2 was met by comparing demographic data of unscreened infants with screened infants. Variables used in the analysis were limited to those variables for which data were immediately available from the two data sets as well as variables that could be derived from the data sets, such as administrative health region and infant mobility.

Objective 3 dealt with only those infants needing follow up. The primary data were those of the newborn screening program. Vital statistics data were used only as a demographic data source to be used for descriptions of the children. Children requiring a repeat screen were determined by the NSP. The NSP data were then searched to see if and when a repeat screen was received. All PKU, Tyrosinemia, and Biotinidase Deficiency repeat screens were done by either the NSP in Edmonton or the Calgary Genetic Diseases Laboratory thus it was possible to determine precisely the degree of follow up for these two disorders. For CH, follow up could be done by a private or public laboratory and in those instances NSP had no record of rescreening.

Section 2: Data Preparation

Data preparation consisted of several major steps. These were:

1. Abstracting the relevant NSP data from a 'lab file' into a working 'screen file';
2. Cleaning and verifying the screen data;
3. Transferring the data from the screen file into a database program;
4. Cleaning and verifying the data in preparation for merging with the Alberta Registries Birth file, the 'birth file';
5. Obtaining data describing all live births in 1992 from Alberta Registries and preparing these data for merging with the screen file; and,
6. Merging the screen file and birth file to make a final file for analysis, the 'analysis file'.

Data for this study came from two major sources: (I) the Registration of Live Births as reported to the Alberta Registries of the Government of Alberta, and (ii) data collected by the Newborn Metabolic Screening Program and maintained by the Department of Laboratory Medicine of the University of Alberta Hospitals. A third source for some repeat data was the Biochemical Genetics Laboratory in Calgary who verified whether specific infants had repeat tyrosine, biotinidase, or phenylalanine samples analysed. In each case, approval for use of these data was obtained. For the Vital Statistics data, final approval was obtained from the Minister of Municipal Affairs. For the screening data, approval was obtained from the Directors of the Clinical Laboratories of the Walter McKenzie Centre in Edmonton and the Biochemical Genetics Laboratory in Calgary, and from the Director of the Newborn Screening Program. Ethics approval for the project was obtained also from the Ethics Committee of the University of Alberta Faculty of Medicine.

Details describing data abstraction from the Alberta Registry and NSP data files are presented in Appendix 1 "Data Abstraction". The data initially obtained from the screen file are listed in Table 2-1. The "gold standard" against which coverage was to be assessed was the data set describing all live births in Alberta in 1992 as provided by Alberta Registries. These data provided the information listed in Table 2-2. The only variables common to both data sets were: the infant's date of birth, the last name of the

Table 2-1. Description of the Variables in the Screen Data File.

Variable Name	Variable Definition
FileId	Name of the raw data file
LineNum	The line number of the raw data file for a particular record
CaseNum	A unique ID number for the record
Gender	The gender of the child
LastName	The last name of the child
FirstName	The first name of the child
ReportAge	The age of the child when the report was filed
AgeType	Days or Months
BloodSource	The geographic source of the blood sample
HospID	The ID number assigned by the source hospital to the child
Accession Number	The accession number created by the WMC laboratory
Doctor	The doctor of record
BloodDate	The date when the blood sample was taken
BloodHour	The time when the blood sample was taken
ReportDate	The date when the NSP archived the screening report
TSHValue	The value of TSH determined by the laboratory
TSHInterp	The interpretation of the TSH value
AminoAcid	The interpretation of the Amino Acid result
Biotin	The interpretation of the Biotinidase result
Comment	A precis of any comment appended to a record
CBD	Is this the child's correct name and birth date (0=No, 1=Yes)
Repeat	Is this a repeat sample, if so, what is the sequence of repeats

infant, the gender, and, to the degree that the place where blood was obtained reflected the infant's place of birth, the place of birth (Table 2-3).

Table 2-2. Description of the Variables in the Birth File

Last Name of Child	Given Name(s) of Child
Sex	Date of Birth
Birth Weight	Gestation
Kind of birth (twin, triplet) and sequence	Hospital of birth
Total number of children ever live born to this mother	Total Number of children ever stillborn to this mother.
Maiden Name of Mother	Mother's Date of Birth
Marital Status of Mother	Mother's AHCIP Number (rare)
Normal Residence of Mother	Postal code of Normal Residence
VID: Unique identifier number (assigned)	

Table 2-3. Available Variables for Linking and for Assessing Missing Infants[**]

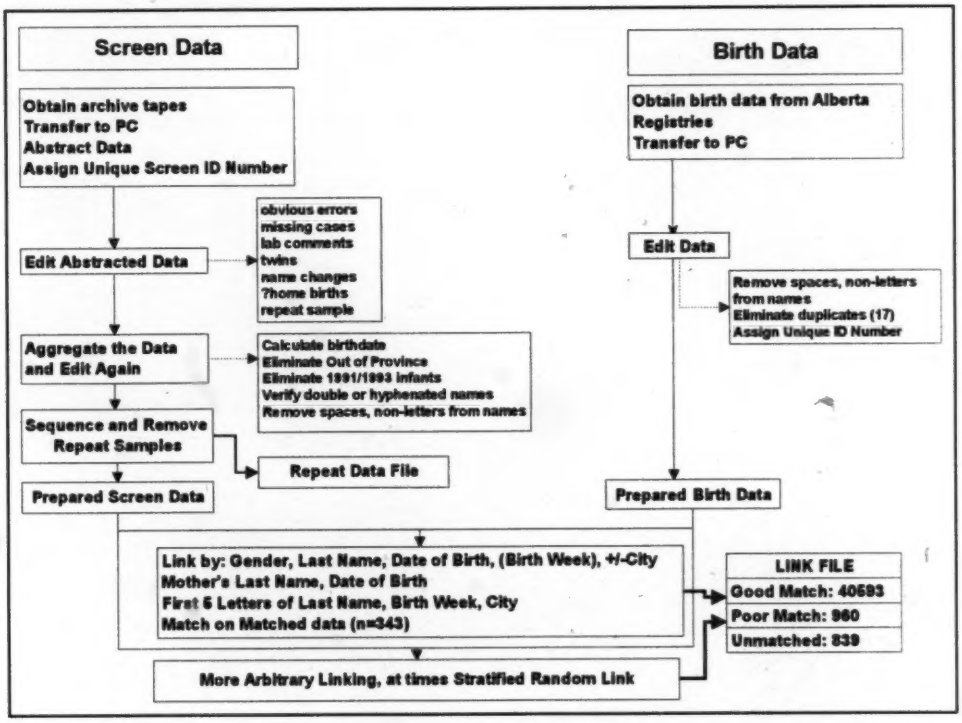
Variables Common to Both Data Sets	Variables Recorded only in Registry of Birth	Variables Recorded only on Neonatal Screening Form
Date of birth	Birth weight	Is baby >24 h old
Last name of child	Normal city of residence	Date of sample
Place where blood was taken (considered to be equivalent to city/hospital of birth)	Postal code of mother's residence	Age at screening
Sex	Mother's age	PKU result
First Name of Child	Mother's marital status	TSH result
	Number of live children born	Biotinidase result
	Time of birth	Tyrosine result
	Kind of birth (twin, triplet)	Time of Sample
	Mother's AHCIP number	
	Last name of mother	
	First name of mother	
	Gestation	

** The variables on the Newborn Screening Form are all supposed to be reported but this is not done consistently.

Section 3: The Match

Details of the matching process are provided in Appendix 2 "The Matching Procedure" and are summarized in Figure 2.1. This process consisted of an iterative series of SPSS MATCH procedures and of hand matching of names and birth dates from the screen file with names and birth dates from the birth data file.

Figure 2-1 The Match Procedure



The result was a master output file consisting only of basic data regarding the screen and birth data files but which could be used ultimately to link these two files with each other. The output file had four variables: (I) the Vital Statistics ID number (VID); (ii) the Screen ID number (SCRID); (iii) the level of confidence ascribed to the match; and (iv) the source of that particular match. With this file it became possible to link data from both the screen data file and the birth data file and to extract only the relevant variables

from either of these two master files.

Each matched record was assigned a level of confidence describing a subjective estimate of the quality of the match for that record. Table 2-4 summarizes the matching criteria, the confidence levels applied, and the numbers matched at each confidence level. An infant was considered to have very good match status [GM] if the confidence of the match was high, and to have uncertain match status [UM] if the confidence of the match was low.

Once the match was done to satisfaction, certain redundant variables were deleted from the various files. Unique personal identifiers, last name, first-name, mother's last name and street address were also deleted from the files. The mother's age at the birth of the child was calculated and then her birth-date was also deleted. Certain new variables were created, relating primarily to the regional distribution of data, the relative frequency of children residing in or being delivered in a particular community, the age at when blood was obtained, and filtering variables to permits rapid data analysis. A variable "Native" was created by considering that communities residing on native reserves were coded as having a high native population. As well, if a community was in a remote area and near a native reserve it also was coded as having a high native population.

Data describing deaths of infants in 1992 were provided by Alberta Health. These data were transferred to a data base but were used during the match only insofar as a child who was going to be arbitrarily matched (a poor match) was not matched against a child who died, unless the death could have occurred after the blood was taken.

Table 2-4. Summary of Matching Criteria, Confidence Levels, and Match Progress

Series	Match Criteria: entered in sequence of match	Confidence Level		Infants Left
		Good	Poor	
0	Initial Birth Registry Data			42392
1	Sex, LastName, FirstName, Birthdate, City	5332		37060
	Sex, LastName, Birthdate, City	28906		8154
	Sex, Mother's LastName, Child's FirstName, Birthdate, City	12		8142
	Sex, Mother's LastName, Birthdate, City	3127		5015
2	Hand Match (spelling, split names, wrong gender, wrong city)	1020		3995
3	LastName, Birthdate (no duplicates in cell)	655		3340
	LastName, Birthweek (no duplicates in cell)	132		3208
	Mother's LastName, Birthdate (no duplicates in cell)		46	3162
	Mother's LastName, Birthweek (no duplicates in cell)		19	3143
	Child's first 5 letters of LastName, Birthweek, city	176		2967
	Mother's first 5 letters of LastName, Birthweek, city		36	2931
4	Hand Match 2	75		2856
5	Match on Previously matched children (343 records removed)			
6	Soundex matches on child's or mothers last name, birthweek, and with or without city agreement		24	2832
	Birthweek and city (to account for screens only)		551	2281
7	Other Hand matches (allocation on basis of geographic and temporal criteria)		248	2033
Subtotal		**39435**	**924**	**2033**
Match Summary for June/October "missing" data*				
1	Sex, Child's Last name or Mother's LastName, Birthdate, City (apparent missing data from June and October)	1087		946
2	Hand matches on June/October "missing data"	71	36	839
	Total	**40593**	**960**	**839**

*These data were collected after examination of the initial match revealed the likelihood of missing data for two periods (one in June and one in October). This match was done later and the results data added. See Appendix 2.

Section 4: Statistical Analysis

Completeness of Coverage

The matched file was divided into three groups: screened infants with a very good match (GM), screened infants with an uncertain match (UM), and unscreened infants with no match (NM); the unscreened infants were those who remained unmatched. The GM and UM groups were combined to obtain an estimate of coverage only. Other analyses compared the groups separately. Figure 2-2 illustrates how the groups were combined to perform the various analyses. The data were divided by (I) region of birth, (ii) the region of blood source, and, (iii) the home region. For most calculations, the birth region of the child was considered the 'region of reference' for statistical and comparative purposes. Two descriptors: less than 24 hour samples and over seven day samples, were derived only from the GM data. The statistical significance of categorical data was assessed using contingency table analysis and Chi-Square statistics and of continuous data using Analysis of Variance with post-hoc comparisons using Scheffe's procedure and a probability criterion of $p<0.05$.

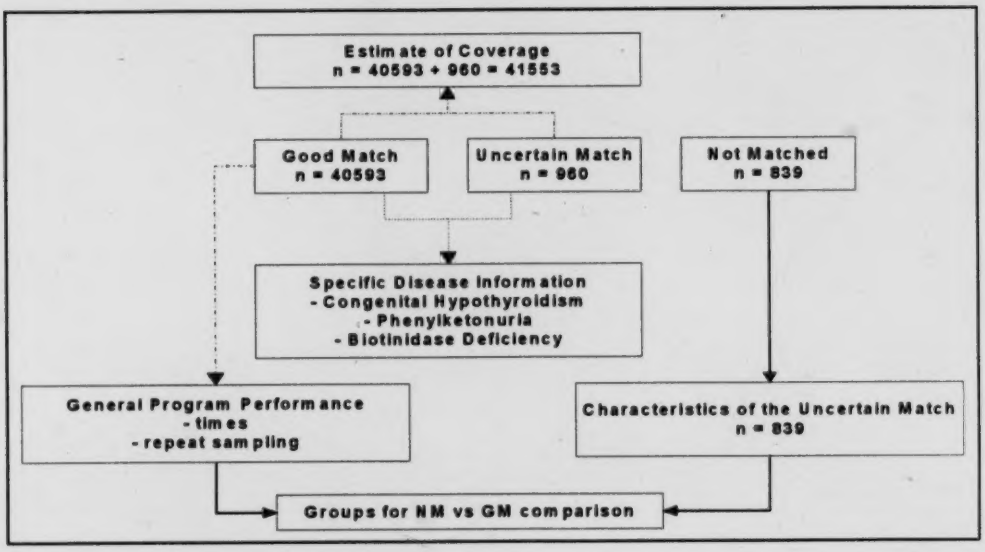


Figure 2.2 Data Combinations for Analysis

Descriptive statistics were obtained also for infants born out of hospital, or in instances where the results suggested an unusual or unexpected pattern of results. The continuous variables of birthweight and maternal age were also divided into categories for use in contingency table analysis. Birthweight was divided into <1.5 kg, 1.5 - 2.5 kg, and >2.5 kg; these are commonly used birthweight categories. Maternal age was divided into <21 years, 21 to 35 years and >35 years; these reflected young, average age and older mothers and also provided adequate sample sizes for analysis.

After initial exploration comparing single variables to matching status, logistic regression was used to control for several factors simultaneously, using matching status (GM vs NM) as the dependent variable. The UM group of infants was not considered in this analysis. Those variables found to be significantly associated with matching status in the single variable analysis above were included as independent variables.

Program Performance in Infants Who Were Screened

For GM infants, descriptors relating to screening program performance were obtained. These related to timing of blood samples, number of unsatisfactory or abnormal results, and timeliness of indicated rescreening. For infants with an abnormal phenylalanine, tyrosine, or biotinidase result, particular effort was made to determine if repeat samples were obtained and to describe the specific program performance for these infants. For abnormal thyroid results, this was not possible, as repeat samples were not necessarily done by the NSP or by the Biochemical Genetics Laboratory in Calgary.

Analysis of 'Irregular Screening Events'

An 'irregular screening event' was defined as an occasion when the child did not receive the recommended standard of screening performance, for example, having an abnormal result that was not repeated. It does not include not being screened at all. Descriptive statistics of infants with screening events were compared to those infants who received recommended care. For infants needing a repeat sample, comparison was with an infant who had a repeat screen requested and done. The variables and methods of comparison were similar to that for the matched/not matched comparisons described earlier except that other variables relating to sampling and results were also used.

CHAPTER 3: RESULTS

Section 1: The Data

Vital Statistics Data

Basic demographic data for the registered as liveborn in Alberta in 1992 are shown in Table 3-1 and are self explanatory. While the total number of infants registered is 42392, the number of cases for individual variables varies slightly because of missing data, most notably for birth weight. Almost 75 percent of infants were born to married women. A further 22 percent were born to women who had never married. The remaining 3 percent were born to women who had formerly been married but were now divorced, widowed, separated, or whose marital status was unknown. For analytic purposes the mothers in these last 4 groups were combined to form one category. There were 1554 infants (3.7%) where the probability of being of native status was high, based on their home residence. In the analysis below, the term "Native" applies to this group. There were 248 deaths; 87 occurred among good match (GM) infants. One hundred sixty seven infants (67%) died in the first week of life, 78 on the first day. Among NM infants, 140 died in the first week of life, 108 within the first 48 hours. Among GM infants, only 13 died in the first week of life. There were 226 infants who were born out of hospital; these were considered as "home births".

The frequency of births in any community varied considerably between regions (Table 3-2). Two communities, Calgary (Region 4) and Edmonton (Region 10), each with over 13000 births per year, accounted for 63 percent of all births with virtually all births occurring in busy obstetrical units with over 1000 births per year. Communities with between 100 and 1000 births per year accounted for about 24 percent of births, and communities with less than 100 births per year accounted for most of the rest. For the 226 infants born out of hospital and where there is no certainty as to the community in which they were born, their birth region was considered to be the same as the region of their home.

Table 3-1. Descriptors of the 1992 Birth Cohort

Variable	Sub Group	Mean	SD	n	% of total births
Total Live Births				42392	100.0
Gender :	Male			21635	51.0
	Female			20752	48.9
	Not stated			5	<0.1
Death 1st Year:	No			42144	99.4
	Yes			248	0.6
Birth Institution:	Hospital			42164	99.5
	Out of Hospital			226	0.5
Multiple Birth:	No			41407	97.7
	Yes			985	2.3
Birth weight (g)		3358	588	42328	
Gestation (was)		39.0	2.3	42389	
Mean Total Liveborn Infants		2.0	1.2	42392	
Mean Total Stillborn Infants		<0.1	0.1	42392	
Mother's Age (y)		27.9	5.4	42386	
Marital Status:	Married			31637	74.6
	Single			9571	22.6
These 3 groups combined to form one group for analysis	Divorced/ Separated			1050	2.5
	Widowed			54	0.1
	Other/Unknown			80	0.2
Probability of Being Native:	Low			40838	96.3
	High			1554	3.7

Table 3-2 Total Births, Stratified by Health Region of Birth

Region	Total Births per Hospital per Year				
	> 1000 births/yr	500 to 999 births/yr	100 to 499 births/yr	< 100 births/yr	Born Out of Hospital
1	1605	0	326	419	2
2	0	895	231	80	2
3	0	0	420	114	5
4	13089	0	0	17	32
5	0	0	298	140	2
6	1923		400	381	11
7	0	0	687	249	5
8	0	0	369	181	8
9	0	0	667	4	12
10	13144	1428	0	39	102
11	0	0	514	71	12
12	0	0	1095	183	10
13	0	954	0	362	7
14	0	0	367	30	1
15	0	0	357	25	3
16	0	696	0	2	1
17	0	0	395	9	6
Unknown	0	0	0		5
Total	**29761**	**3973**	**6126**	**2306**	**226**

The Screen Data

There were 43155 records of screen data; 43125 were obtained from the WMC site and a further 30 records were obtained from Alberta Children's Hospital. Initial screen results accounted for 41553 records and 1602 were repeat screen records from 1485 infants (Table 3-3). During the match procedure repeat screen records for 63 infants were found but for which there was no apparent matching initial screening record; 49 were in the GM group. There were 1412 initial records that required a repeat because of an abnormal result, a poor sample, or samples taken too early in life.

Table 3-3 Source and Number of Screen Records

Descriptor	n
Total Screen Records Found	43155
From WMC Records	43125
From Calgary Records	30
Order of Records	
Initial Record	41553
Records Considered Repeat	1602
Records Repeated For Cause	668
Repeat Records WITHOUT Initial Record	63
Status of Record	
Abnormal Initial Record	1177
Blood Obtained Before 24h	235

Section 2: Coverage

In the 42392 infants in the Vital Statistics database, 40593 were in the GM group. Of the remaining 1799 infants, 960 were used to match the remaining unmatched records and 839 infants remained unmatched. The sum of 41553 matched infants is the most likely estimate of coverage and is 98.0 percent of the eligible population.

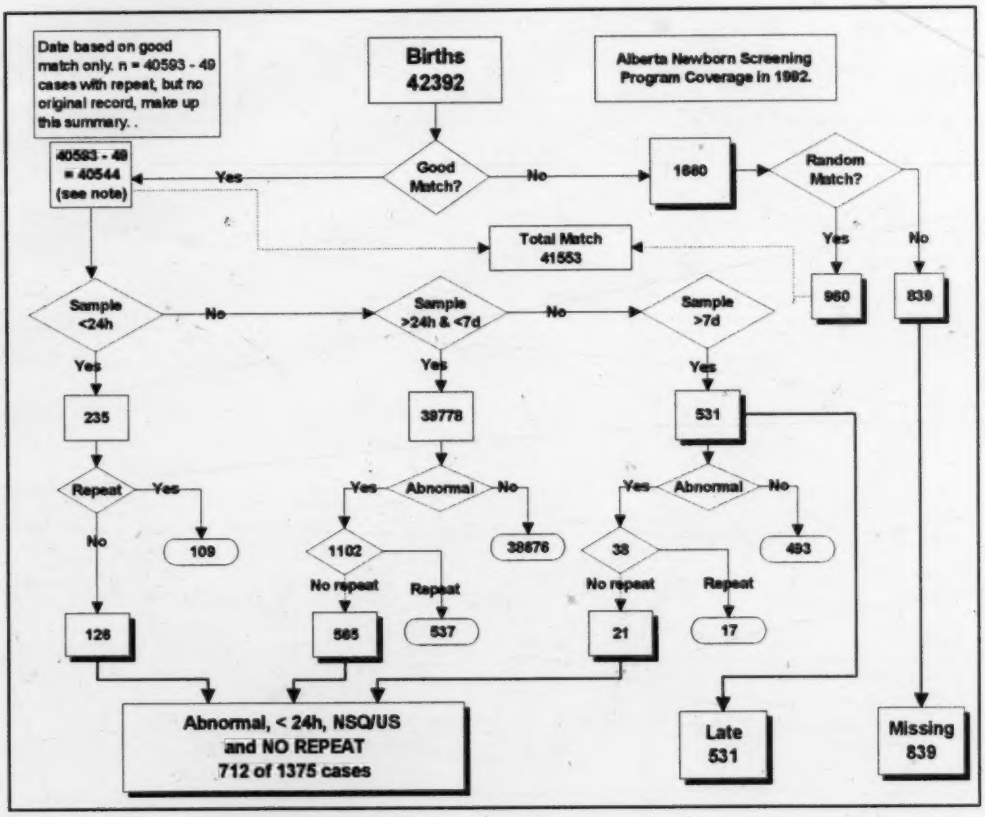


Figure 3-1 A Summary of the Results

A graphical accounting of the children is shown in Figure 3-1, which accounts for all the children but provides details primarily for the GM group. The key features of Figure 3-1 are that: (a) among GM infants there were 1375 cases where a repeat sample was requested; (b) of these 1375 cases, in only 662 was a repeat sample found; (c) there were 531 children who were sampled initially after seven days of age; and (d) there were 839 more children born than were screened.

Variation Among Health Regions

Table 3-4 provides a summary of some key variables for the entire data set and also provides data about early and late blood samples for only GM infants. Data are divided by the (I) region of birth, (ii) region of blood source, and (iii) home region. Unless otherwise stated, the birth region of the child is considered the 'region of reference' for statistical and comparative purposes.

Except for infants whose home region was out of province the proportion of GM infants in any region ranged from 93.5 percent to 97.8 percent. There was considerable mobility of infants from the community of birth to the home community, with the less populated communities having significant numbers of infants born in cities with large hospitals. This was particularly evident for Edmonton, which had about 3900 more births than would be expected on the basis of residence.

Infants Born Out of Hospital

There were 226 infants born out of hospital (OOH) and considered to be 'home' births. Of these only 167 (73.9%) were matched. The proportion in any home region varied from 0.09 percent in Region 1 to 1.2 percent in Region 17, although in each instance the absolute numbers were very small, the highest number (102) being born in Region 10 and the next highest number (32) in region 4. Infants born out of hospital had higher mean birth weight and their mother's were on average older and more likely to have had other children (Table 3-5).

Characteristics of the Good Match, Uncertain Match, and Not Matched Groups

Data describing the three groups of infants are found in Tables 3-6 through 3-9. Compared to the GM group, UM infants more often died in infancy, more commonly were of low birth weight and were born after a shorter gestation (Table 3-6). Nearly 5.4 percent of NM infants were born out of hospital, compared to only 0.4 percent of GM infants. As well, NM infants were slightly more likely than were GM infants to have come from a community with a high native population, and were more likely to have moved, particularly from the region in which they were born (Table 3-7).

Table 3-4 A Numerical Description of Key Variables

Data Descriptors for Entire Data Set										"Good Match" Only	
Region	Blood Site n (priv lab)	Birth Region n	Blood Source n	Home Region n	Out of Hospital n	Not Matched n (% birth)	Uncertain Match n (%birth)	Good Match n (% birth)	Need Repeat n (% repeat)	Blood <24 h n	Blood >7 d n
1	12 (1)	2352	2296	2285	2	50 (2.1)	40 (1.7)	2262 (96.2)	93 (47.3)	8	41
2	6 (1)	1208	1185	1196	2	18 (1.5)	23 (1.9)	1167 (96.6)	38 (39.5)	10	30
3	6	539	527	911	5	7 (1.3)	10 (1.9)	522 (96.8)	14 (35.7)	3	31
4	9 (4)	13138	12912	12313	32	257 (2.0)	328 (2.5)	12553 (95.5)	467 (54.8)	101	109
5	5	440	428	782	2	6 (1.4)	4 (0.9)	430 (97.7)	9 (66.7)	1	7
6	13	2715	2776	2646	11	37 (1.4)	66 (2.4)	2612 (96.2)	70 (64.3)	14	28
7	16	941	913	1042	5	14 (1.5)	7 (0.7)	920 (97.8)	30 (66.7)	5	15
8	6	558	534	1361	8	6 (1.1)	10 (1.8)	542 (97.1)	18 (66.7)	6	9
9	4	683	552	1456	12	12 (1.8)	14 (2.0)	657 (96.2)	38 (50.0)	5	8
10	11 (6)	14713	14520	10801	102	315 (2.1)	359 (2.4)	14039 (95.4)	474 (39.5)	55	186
11	8	597	557	1183	12	7 (1.2)	16 (2.7)	574 (96.1)	17 (70.6)	2	23
12	14	1268	1239	2565	10	23 (1.8)	12 (0.9)	1253 (97.3)	27 (37.0)	5	14
13	8	1323	1282	1326	7	29 (2.2)	29 (2.2)	1265 (95.6)	30 (50.0)	5	8
14	3	398	394	348	1	7 (1.8)	9 (2.3)	382 (96.0)	13 (15.4)	2	7
15	3	385	373	559	3	15 (3.9)	10 (2.6)	360 (93.5)	20 (25.0)	2	4
16	1	699	675	672	1	19 (2.7)	16 (2.3)	664 (95.0)	20 (0.0)	1	0
17	2	410	388	496	6	14 (3.4)	7 (1.7)	389 (94.9)	34 (70.6)	10	10
Out of Prov				446							
Unknown		5		4	5	3 (60.0)		2 (--)		0	1
Total	127	42392	41553	42392	226	839 (2.0)	960 (2.3)	40592 (95.8)	1053(48.4)	235	531

Table 3-5 Characteristics of Infants Born Out of Hospital

Characteristic	Home			Hospital			p
	Mean	SD	n	Mean	SD	n[1]	
Birth weight	3459	702	224	3358	587	42104	0.01
Gestation	39.1	3.24	226	39.0	2.18	42163	0.84
Total number of liveborn	2.69	1.76	226	2.00	1.17	42166	0.12
Mother's Age	30.7	5.3	226	28.0	5.4	42160	0.00
Age when Initial Blood Obtained	6.65	5.8	167[2]	2.56	1.93	40376	0.00

1. 'n' varies because of missing data

2. Only GM infants used.

Birthweight

There is a strong association of birth weight with matching status, the low birth weight infant was much more likely to be unmatched or poorly matched than was an infant whose birthweight was over 2500 g. Overall, 6.2 percent of infants had a birthweight of less than 2400 g and 1.2 percent had a birthweight of less than 1500 g; however, in the NM group, 32.1 percent, and in the UM group, 14.9 percent of infants were under 2500 g, but only 5.5 percent of the GM group were under 2500 g (Table 3-8a). For infants with a birthweight of under 1500 g, the differences between the NM and GM groups are even more striking (18.6% vs 0.7%). Because so many NM infants died in infancy, birth weights excluding deaths in the first week of life were compared (Table 3-8b). Even when first week deaths are excluded, infants under 2500 g constituted 24.2 percent of NM infants compared to 5.5 percent for the GM group.

Maternal Marital Status

Only 50.8 percent of NM infants had mothers who were married, compared to 76.0 percent of GM infants (Table 3-9). As well, the infants of divorced/separated mothers were much more likely to be NM than were those whose mothers marital status was recorded as single.

Table 3-6 Demographic Characteristics Stratified by Match: Continuous Data

Variable	Not Matched (NM)			Uncertain Match (UM)			Good Match (GM)			Post Hoc Contrasts (Scheffe) p<0.05
	Mean	SD	n	Mean	SD	n	Mean	SD	n	
Birthweight	2740	1129	781	3164	780	954	3375	560	40593	NM<UM<GM
Gestation	35.2	6.5	839	38.2	3.7	959	39.1	1.9	40591	NM<UM<GM
Total Liveborn	2.1	1.6	839	2.1	1.4	960	2.0	1.2	40593	NM > GM
Total Stillborn	0.2	0.4	839	0.1	0.3	960	<0.0	0.02	40593	NM>UM>GM
Mother's Age	27.4	5.9	838	27.2	6.3	958	28.0	5.3	40590	NM,UM<GM
Age at Death	3.4	18.0	146	4.7	5.4	16	72.3	69.1	86	NM,UM<GM

Table 3-7 Demographic Characteristics of the NM, GM, and UM Groups: Place of Birth and Mobility

Variable	Not Matched		Uncertain match		Good Match		Total		Chi-Square	
	n	%	n	%	n	%	n	%	x^2	p
Hospital Birth										
Yes	794	94.6	947	98.6	40425	99.6	42166	99.5	392	<0.05
No	45	5.4	13	1.4	168	0.4	226	0.5		
Native Community										
Yes	52	6.2	49	5.1	1453	3.6	1554	3.7	21.7	<0.05
No	787	93.8	911	94.9	39140	96.4	40838	96.3		
Mobility										
Did Not Move	543	64.7	718	74.8	19812	73.4	31073	73.3	52.5	<0.05
Moved in Region	134	16.0	93	9.7	5520	13.6	5747	13.6		
Moved Out of Region	162	19.3	149	15.5	5261	13.0	5572	13.1		

Table 3-8a Frequencies of Birthweight in All Children

Birth Weight (g)	Not Matched		Uncertain Match		Good Match		Total	
	n	%	n	%	n	%	n	%
<1500	145	18.6	48	5.0	292	.7	485	1.1
1501 - 2500	106	13.6	94	9.9	1956	4.8	2156	5.1
2501 - 6000	530	67.9	812	85.1	38345	94.5	39687	93.8
Total	781		954		40593		42328*	
Chi-Square	2482		p<0.05					

* n=42328 because of missing data for birthweight

Table 3-8b
Frequencies of Birthweight in Children Alive at Age 1 or Dying After Age 7 Days

Birth Weight (g)	Not Matched		Uncertain Match		Good Match		Total	
	n	%	n	%	n	%	n	%
<1500	74	11.4	44	4.7	288	.7	406	.9
1501 - 2500	83	12.8	91	9.7	1956	4.8	2130	5.1
2501 - 6000	491	75.8	805	85.6	38335	94.5	39631	94.0
Total	648		940		40579		42167	
Chi-Square	1080		p<0.05					

Table 3-9 Match Status by Marital Status of the Mother

Marital Status	Not Matched		Uncertain Match		Good Match		Total	
	n	%	n	%	n	%	n	%
Married	426	50.8	375	39.1	30836	76.0	31637	74.6
Single	294	35.0	362	37.7	8915	22.0	9571	22.6
Divorced	107	12.8	208	21.7	735	1.8	1050	2.5
Widowed	6	0.7	10	1.0	38	0.1	54	0.1
Unknown /Other	6	0.7	5	0.5	69	0.2	80	0.2
Total	839		960		40593		42392	
Chi-Square	Group	Value	p	Group[1]	Value	p		
	All	2364	<0.05	SvsDWUO	666	<0.05		

1. Comparison is Single vs Divorced, Widowed, Unknown/Other

Maternal Age

Table 3-6 showed that the mean maternal age was greater in the GM compared to the UM group. Table 3-10 summarizes the relationship of three categories of maternal age with matching status and shows that mothers aged less than 21 years are disproportionately represented in the NM and UM groups. Of the 839 women in the NM group, 14.5 percent were under 21 years, compared to 9.1 percent of the GM group; among the 960 mothers of UM infants, 18.2 percent were under 21 years of age. Older mothers were fairly evenly distributed between the two groups (9.4% and 8% respectively).

Table 3-10 Mother's Age by Age Category and Match Status

Age (y)	Not Matched		Uncertain Match		Good Match		Total	
	n	%	n	%	n	%	n	%
<21	122	14.5	175	18.2	3693	9.1	3990	9.4
21 - 35	645	76.9	687	71.6	33669	82.9	35001	82.6
35+	72	8.6	98	10.2	3231	8.0	3401	8.0
Chi-Square	131	p <0.00						

Relationships Between Key Variables and Match Status

Table 3-11 presents summary data describing matching status as a function of maternal age, birthweight and marital status. Data for infants whose birthweight was missing are excluded. With the exception of the "other" less than 21 year group, where there were only 20 cases, matching was most likely to occur among infants with a birth weight of over 2500 g born to married mothers aged 21 to 35 years. In this group 98.3 percent of infants were in the GM group. At every age, as birthweight rises, the proportion of infants matched also rises. The proportion of GM infants rises also as the marital status changes from "other" to "single" to "married". Only about 71 percent of infants whose mother's marital status was "other" were GM whereas over 90 percent of

Table 3-11 Marital Status, Birthweight, and Age Relationships[1]

Marital Status	Age (y)	Weight (g)	No Match		Uncertain		Good Match		Total	Percent GM
			n	%	n	%	n	%		
Single	< 21	<2500	31	27.4	17	10.6	177	5.9	225	78.7
		2500+	82	72.6	143	89.4	2828	94.1	3053	92.6
	Total for Group		**113**	**3.4**	**160**	**4.9**	**3005**	**91.7**	**3278**	
	21-35	<2500	47	29.2	23	2.0	404	7.1	474	85.2
		2500+	114	70.8	169	88.0	5277	92.9	5560	92.1
	Total for Group		**161**	**2.7**	**192**	**3.2**	**5681**	**94.1**	**6034**	
	35+	<2500	4	44.4	5	55.6	22	9.6	31	71.0
		2500+	5	55.6	4	44.4	207	90.4	216	95.6
	Total for Group		**9**	**3.6**	**9**	**3.6**	**229**	**92.7**	**247**	
Married	< 21	<	2	40.0	1	6.7	35	5.2	38	92.1
		2500+	3	60.0	14	93.3	633	94.8	650	97.4
	Total for Group		**5**	**0.7**	**15**	**2.2**	**668**	**97.1**	**688**	
	21-35	<2500	139	40.5	60	19.3	1391	5.1	1590	87.5
		2500+	204	59.5	251	80.7	25913	94.9	26368	98.3
	Total for Group		**343**	**1.2**	**311**	**1.1**	**27304**	**97.7**	**27958**	
	35+	<2500	13	37.1	13	28.9	145	5.1	171	84.8
		2500+	22	62.9	32	71.1	2719	94.9	2773	98.1
	Total for Group		**35**	**1.2**	**45**	**1.5**	**2864**	**97.3**	**2944**	
Other*	<21	<2500	0	0.0	0	0.0	4	20.0	4	100.0
		2500	0	0.0	0	0.0	16	80.0	16	100.0
	Total for Group		**0**	**0.0**	**0**	**0.0**	**20**	**100.**	**20**	
	21-35	<2500	11	11.5	16	8.8	59	8.6	86	68.6
		2500+	85	88.5	165	91.2	625	91.4	875	71.4
	Total for Group		**96**	**10.0**	**181**	**18.8**	**684**	**71.2**	**961**	
	35+	<1500	4	21.1	7	17.1	11	8.0	22	50
		2500+	15	78.9	34	84.5	127	92.0	176	72.2
	Total for Group		**19**	**9.6**	**41**	**20.7**	**138**	**69.7**	**198**	

1. Figures in the highlighted column are to be read and interpreted across the page; figures not highlighted are to be read down the page. The percent GM is calculated from the raw GM data and the total in the respective rows.

infants of single and married mothers were GM. The poor GM match in the "other" group appears to be unrelated to either maternal age or infant birthweight. These data suggest that match status is a function of the combination of maternal age, marital status, and birthweight. These potential relationships were explored further with logistic regression.

Logistic Regression Analysis of Matched vs Not Matched Infants

Logistic regression analysis was done with matching status as the dependent variable and with an inclusion probability of 0.05. Initial exploration of the data using logistic regression eliminated the variables describing probability of native status and birth region as important determinants of match status. Variables that were included in the final regression were death in the first week of life (2 categories), birth weight (3 categories), marital status (3 categories), and birth frequency (5 categories). The results of this regression are generally consistent with the stratified analysis done previously (Table 3-12) except that the mother's age and infant mobility were not significant predictors of non-matching.. Infants who died in the first week of life were 400 times as likely to not be matched than were those who did not die, or who died after age seven days. Infants with a birth weight of under 1500 g were nearly 19 times as likely to be unmatched than were infants with a birthweight over 2500 g. Infants with a birth weight between 1500 - 2499 g were also 3.3 times more likely to be NM. Infants of mothers living alone were at greater risk of being not matched than were infants of mothers who are married, however, infants of mothers who had never married (OR: 2.6) were at less risk than infants of mothers who were widowed, separated, or divorced (OR: 12). When the child was born out of hospital, the probability of not being matched was nearly 20 times that of infants born in large hospitals.

In summary, 98 percent of infants were screened; however, this implies that 839 infants were not screened. Of these 839, those at relatively greater risk of not being screened include those who were born out of hospital, or were of low birth weight, born of unmarried or formerly married mothers, or who were sufficiently sick as to be at high risk of death.

Table 3-12 Logistic Regression: Not Matched Compared to Good Match

Variable	Code	B	OR	95% CI	p
Birth Weight (g)					
>2500	REF	0	1		
1500-2499	BW2	1.181	3.26	2.56 4.14	<0.05
<1500	BW1	2.964	19.38	14.46 25.98	<0.05
Marital Status					
Married	REF	0	1		
Single	MS1	0.974	2.65	2.22 3.16	<0.05
Formerly married	MS2	2.549	12.80	10.06 16.29	<0.05
Births per year					
>1000	REF	0	1		
500 - 999	BF1	0.154	1.17	0.89 1.53	ns
100 to 499	BF2	0.152	1.16	0.92 1.47	ns
< 100	BF3	-0.168	0.83	0.54 1.27	ns
Born Out of Hospital	BF4	2.987	19.81	13.44 29.22	<0.05
Died in First Week (0:No, 1:Yes)	D1W	5.997	400	216 742	<0.05
Constant		-4.928			<0.05

No match status = 1; Good Match Status = 0

Section 3: Program Performance

The above has provided a description of who were screened and who were not. The following provides data on program performance and relates only to those who were screened.

Blood Sources

The 41553 initial samples were received from 127 different hospital or laboratory sites. All except 1119 samples were received from a hospital or hospital laboratory; the 1119 came from one of 17 private laboratories, principally in Edmonton or Calgary. There was wide variation in the frequency of collection of blood among collecting sites. Almost 68 percent of samples were received from ten collecting sites, principally in Calgary and Edmonton, whereas 100 other sites accounted for only 15 percent of samples. Fifty four sites sent fewer than one sample per week.

Timing of Initial Blood Samples

Of the 40543 GM infants, 98.1 percent had samples taken at the desired time of one to seven days of age; 0.58 percent (235) had samples taken before 24 hours, and 1.3 percent (531) had samples taken after seven days (Table 3-13).

Table 3-13 Characteristics of Infants with Blood Taken When Over 7 Days of Age

Variable	Age When Initial Blood Obtained						Post-hoc Contrast (Scheffe) $p < 0.05$
	<24h (1)		1 - 7 Days (2)		> 7 Days (3)		
	n = 235		n = 39777		n = 531		
	Mean	SD	Mean	SD	Mean	SD	
Birthweight (g)	3350	696	3381	549	2913	969	1,2>3
Gestation (wk)	38.7	2.8	39.1	1.8	36.9	4.4	2>1>3
Total Liveborn Infants	2.0	1.1	2.0	1.2	2.2	1.4	3>1,2
Age (y) Mother	28.4	5.3	28.0	5.3	28.0	5.4	
Age (d) Infant	0	0	2.5	0.9	13.2	11.3	3>2

The mean age of blood sampling was 2.6 days overall. For infants with blood obtained between one and seven days, the mean age was 2.5 days; for those with initial blood samples obtained after seven days, the mean age was 13.2 days.

About 8.5 percent of the infants sampled at less than 24 hours had birth weights of under 2500 g, compared to 5.2 percent for infants sampled between one and seven days (Tables 3-13 and 3-14). In contrast, infants who had their blood sample taken at over seven days of age, when compared to those whose blood was taken between one and seven days of age, were much more likely to be of lighter birthweight and earlier gestation, were less likely to be born to a married mother, were slightly more likely to be born in a community where less than 100 infants were born in a year, were more likely to have been born in a community different than their normal residence, and were much more likely to have been born out of hospital. Of the 167 infants born out of hospital for whom age at sampling can be calculated, 61 (36.5%) were sampled when over seven days of age, whereas only 1.3 percent of all infants were first sampled this late in life. Infants born out of hospital had their initial blood sample obtained at an average age of 6.65 days.

Twenty percent of infants with a birthweight of less than 1500 g were sampled at over seven days as compared to only one percent of infants with birth weights of over 2500 g. Overall, 29.6 percent of infants sampled at over seven days had birth weights under 2500 g, compared to five percent of infants sampled earlier in life.

The combination of late sampling in a high proportion of low birth weight infants and infants born in smaller communities suggests that perhaps low birth weight infants were being born in small communities and transferred to larger institutions for care and the screen was delayed. This is not the case. Of the 95 infants who "moved" between regions (i.e. their birth town was in a region different from their home town), 49 were less than 2500 g; of these, 46 were born in large hospitals in Edmonton or Calgary. Fifteen infants with birth weights less than 2500 g "moved" within a region. Ten of these were born in medium sized communities and five were born in small communities. Among the 285 infants who did not move and yet had late samples, 93 had a birth weight of less than 2500 g and 82 of these infants were born in busy hospitals (data not shown in tables).

Table 3-14 Comparison of Infants by Timing of Blood

Variable	Age When Blood Taken						Chi-Square
	< 24 h		1 to 7 Days		Over 7 Days		
	n	%	n	%	n	%	
Births per Year							
>1000	165	70.2	27955	70.3	281	52.9	
500 - 999	17	7.2	3765	9.5	40	7.5	149.6 p<0.05
100 - 499	37	15.7	5786	14.5	98	18.5	
<100	14	6.0	2168	5.5	51	9.6	
Out of Hospital	2	0.9	104	0.3	61	11.5	
Birthweight (g)							
<1500	8	3.4	223	0.6	59	11.1	516 p<0.05
1500-2499	12	5.1	1840	4.6	98	18.5	
>2500	215	91.5	37715	94.8	374	70.4	
Death							
No	235	100	39697	99.8	525	98.9	9.87[1] p<0.05
Yes	0	--	81	0.2	6	1.1	
Marital Status							
Married	175	74.5	30244	76.1	380	71.6	8.36 p<0.05
Single	59	25.1	8719	21.9	125	23.5	
Wid/Div/Sep/Oth	1	0.4	815	2.0	26	4.9	
Moved							
No	171	72.8	29322	73.7	288	54.2	39.2 p<0.05
Moved in Region	30	12.8	5333	13.4	149	28.1	
Moved out of Region	34	14.4	5122	12.9	94	17.7	
Total	**235**		**39778**		**531**		**40544**

[1] Because of the empty cell, a likelihood ratio chi-square was calculated for these data, with the 0 replaced by 0.00001 to permit computation.

These results suggest that there were two groups of infants who made up the late sampled group. There was a group of low birth weight infants who were high risk pregnancies in small communities and were referred to major hospitals before delivery and the child's postnatal medical status precluded sampling at the desired age and there was a second group of normal weight infants who were not sampled at the appropriate time.

Blood Samples Requiring Repeat

There were 1412 records requiring a repeat sample because the result was abnormal or they were taken before 24 hours, or were of unsatisfactory quality. Of these, 1375 were from GM infants. Reasons for the need for repeat sampling are shown in Table 3-15 as well as the proportion of infants for whom a repeat sample was found.

Table 3-15 Reasons For the Need to Repeat a Screen Sample

Reason	Good Match	Uncertain Match	Total	Percent Repeated	<24 hour samples
<24 h	235	---	235	46.4	
NSQ/US	140	2	142	65.5	2
Abnormal /Decreased Amino Acid	149	15	164	50.0	1
High Phenylalanine	8	--	8	100.0	--
High Tyrosine	713	18	731	44.2	3
Low Biotin	72	1	73	53.4	1
High TSH	58	1	59	24.1	10
Total	1375	37	1412	47.9	17

Abnormal biochemical results accounted for 1035 records. The most common reason for a repeat was a high tyrosine, accounting for 731 records (51.8%). The other three diseases of interest accounted for about 9.8 percent of abnormal results. Unsatisfactory samples accounted for 139 records. This is about 0.34 percent of all samples, but the proportion varied considerably across regions. Blood samples were

taken before 24 hours of age from 235 infants. Among these infants, there were 17 abnormal biochemical results; this is 7.2 percent of the records obtained and is about 2.5 times the proportion of abnormal results seen in infants initially sampled after 24 hours of age. The most common abnormality was a high TSH, which occurred 30 times more commonly among these infants than among infants whose blood was taken after 24 hours.

Table 3-16 Characteristics of Good Match Infants Who Require Repeat Screen

Variable	Repeat Not Needed			Repeat Needed			
	Mean	SD	N	Mean	SD	N	p
Age (y) Mother	28.0	5.3	39215	27.8	5.6	1375	ns
Birthweight (g)	3384	550	39218	3089	726	1375	<0.05
Gestation (wk)	39.1	1.8	39216	37.8	2.9	1375	<0.05
Total Liveborn (n)	2.0	1.6	39219	2.2	1.2	1375	<0.05
Age (d) when blood taken	2.6	1.9	39170	2.6	2.9	1375	ns
Age (d) Died	75.5	71	77	44.8	44	9	ns

Infants in the GM group needing a repeat sample were more likely to be of lower birthweight and earlier gestation than were those not requiring a repeat sample (Tables 3-16). They were more likely to have a residence different from the community in which they were born (Table 3-17). Nearly four percent of those whose birth town was different from their home town required rescreening whereas only 0.2 percent of infants who did not move required rescreening. Infants whose birth and home town were different had a significantly higher frequency of NSQ/US samples than did infants whose birth and home towns were the same. Infants who were likely to be native also were more likely to need rescreening. Almost 7.6 percent of native infants required rescreening compared to 3.4 percent of all other infants. .

Table 3-17 Characteristics of Infants Needing Repeat Screen: Categorical Data

Variable	Repeat Not Needed	%	Repeat Needed	%	Chi-Square
BirthWeight (g)					
<1500	245	97.2	47	2.8	465 p<0.05
1500 - 2499	1752	89.6	204	10.4	
>2499	37222	97.1	1123	2.9	
Marital Status					
Married	29981	96.9	955	3.1	32.67 p<0.05
Single	8530	95.7	385	4.3	
Other	808	96.0	34	4.0	
Birth Frequency					
>10 births/day	24819	96.6	874	3.4	13 p<0.05
1 - 10 births/day	6827	97.2	198	2.8	
< 1 birth/day	7414	96.2	293	3.8	
Out of Hospital	159	94.6	9	5.4	
Moved					
No	28879	96.9	929	3.1	27.5 p<0.05
Moved IN Region	5149	96.2	205	3.8	
Moved OUT of Region	5191	95.6	240	4.4	
Native					
No	37871	96.8	1269	3.2	68 p<0.05
Probable	1348	92.8	105	7.2	
Dead					
No	39141	96.6	1365	3.4	12.9 p<0.05
Yes	78	89.7	9	10.3	

Logistic regression was used to explore the potential factors to account for the need to repeat. Birth weight, death in the first week of life, and native status were the most important predictors for the need to repeat a sample, with odds ratios ranging from 2 to 5 (Table 3-18). Maternal status, and infant mobility also were significant predictors but, with OR's much closer to 1.0, were of marginal utility.

Table 3-18 Logistic Regression: Infants Needing Repeat vs Those Not Needing Repeat. Good Match Infants Only

Variable	Code	B	OR	95% CI	p
Birth Weight (g)					
>2499	REF	0	1		
1500 - 2499	BW2	1.314	3.72	3.19, 4.35	<0.05
<1500	BW1	1.611	5.01	3.64, 6.89	<0.05
Marital Status					
Married	REF	0	1		
Single	MS1	0.189	1.21	1.07, 1.37	<0.05
Formerly Married	MS2	0.289	1.33	0.99, 1.80	ns
Did Child "Move"					
No	REF	0	1		
Moved In Region	M1	0.097	1.10	0.94, 1.29	ns
Moved Out of Region	M2	0.153	1.16	1.00, 1.36	<0.05
Died in Week 1	DW1	1.205	3.34	1.20, 9.30	<0.05
Native	N1	0.701	2.01	1.62, 2.51	<0.05
Constant		-3.624			

Need Repeat = 1; Do Not Need Repeat = 0

The Non-Repeaters

Of the 1412 infants who had a sample result that required rescreening, in only 668 (47.3%) was a repeat sample found (Table 3-15); 655 of these were infants in the GM group. Compliance was best for infants with high phenylalanine screens (100%) and worst for infants with high TSH results (24.1%). Table 3-19 shows that the 668 infants who were rescreened, when compared to those that were not rescreened, were similar in most available descriptors.

Table 3-19 Good Match Infants Who Had Repeat Screens Compared to Those Who Did Not Have Repeat Screens

Variable	Repeated n = 655		Not Repeated n = 719		p
	Mean	SD	Mean	SD	
Age (d) when blood first taken	2.6	2.1	2.7	3.6	ns
Birth weight (g)	3045	760	3130	691	<0.05
Gestation (wk)	37.6	3.2	37.9	2.7	<0.05
Total infants liveborn to mother (n)	2.1	1.2	2.2	1.2	ns

Infants with birth weighs between 1500 g and 2500 g were less likely to have a repeat sample when requested (41.8%) than were either lighter (50.8%) or heavier (47.1%) infants (Table 3-20). Infants of single mothers also were less likely to have a repeat (41.8%) compared to married mothers (50.8%) or separated, widowed, or divorced mothers (47.1%). Logistic regression comparing the two groups confirms the above observations and extends them. Infants of single mothers were found to be significantly more likely than married mothers not to get the required repeat (OR: 1.51). As well, boys were more likely to get the required repeat than were girls (OR: 0.72) and infants born in hospitals where there were fewer than 500 births per year were more likely to have the required repeat than were infants born in busier hospitals. (Table 3-21).

Table 3-20 Good Match Infants With Repeat Screens Compared to Infants Without Repeat Screens: Categorical Data

Variable	Repeated	Not Repeated	Chi-Square
Birth Weight (g)			6.65 p<0.05
< 1500	33	14	
1500 - 2500	101	103	
> 2500	529	595	
Marital Status			6.96 p<0.05
Married	486	470	
Single	161	224	
Other	16	18	
Births per Year			5.74 p<0.05
> 1000	441	499	
500 - 999	49	72	
100 - 499	123	105	
< 100	44	33	
Born Out of Hospital	6	3	
Moved			ns
No	442	488	
Moved in Region	110	100	
Moved out of Region	111	124	
Native			ns
No	619	651	
Probable	44	61	
Dead			8.4 p<0.05
No	663	703	
Yes	0	9	

Table 3-21
Logistic Regression: Infants Not Having Required Repeat vs Those Having Required Repeat

Variable	Code	B	OR	95% CI	p
Birth Weight (g)					
>2499	REF	0	1		
1500 - 2499	BW2	-0.224	0.80	0.59, 1.08	ns
<1500	BW1	-1.075	0.34	0.18, 0.89	<0.05
Marital Status					
Married	REF	0	1		
Single	MS1	0.41-	1.51	1.18, 1.92	<0.05
Formerly Married	MS2	0.517	1.68	0.92, 3.05	ns
Births per Year					
>1000	REF	0	1		
500 - 999	B1	0.106	1.11	0.75, 1.64	ns
100 - 499	B2	-0.439	0.65	0.47, 0.87	<0.05
< 100	B3	-0.596	0.55	0.34, 0.88	<0.05
Born out of hospital	B4	-1.145	0.32	0.08, 1.31	ns
Gender (M = 1)	G1	-0.324	0.72	0.58, 0.89	<0.05
Constant		0.295			<0.05

Obtained required repeat = 0; Did not obtain required repeat = 1

The Repeaters

There were 655 GM infants who had the required repeat sample taken. The average age for a repeat sample was 27.2 ± 22.4 days, with a range of ages from 1 to 270 days; 65 percent of the samples were obtained by 30 days of age.

Section 4: Specific Diseases

Phenylketonuria

Eight children had high initial serum phenylalanine levels and repeat records were found for all eight children. All children with suspect PKU results had repeat screens at an average age of 11.6 days and a maximum age of 23 days. One other child initially had an abnormal amino acid report which on repeat was shown to be high phenylalanine. Of these nine children, two were verified to have phenylketonuria and the others considered to have hyperphenylalaninemia. One of these with hyperphenylalaninemia had two repeat samples taken, the first at 12 days of age. Both results showed a high phenylalanine but the child was not referred to a pediatrician and was not seen by a dietitian. In fact, the NSP was unaware of this situation until the researcher contacted the Biochemical Genetics Laboratory in Calgary in an effort to track the child down. Inquiries were made and it was realized that the child had not been followed up.

Biotinidase Deficiency

Initial biotin was reported as decreased (DE) in 73 cases and one child was shown ultimately to have biotinidase deficiency. Repeat samples were found for one UM and 39 GM infants. Repeats were taken at an average age of 34.7±18.6 days (range 3 - 92 d) (GM data only). Only 46 percent of infants had a repeat done by 30 days of age. Of the 40 repeats, 35 were considered normal, four were decreased (DE) and one was NSQ. The four DE children were repeated after an average of 21.7 days, and all four came back as DE again, but there was no evidence of a third repeat sample. Three children had a DE Biotin on a repeat sample, but a normal original Biotin. One child was repeated because of a poor initial sample and two children were repeated because of high tyrosine values on the initial samples. All three had the abnormality noted on the second sample and a request for a repeat sample, but there was no record found of any later samples.

Sixteen of the 76 DE biotin results came from one area in Region 1; of these 13 have a record of being repeated. There were 1024 children born who live in this region. Thus the simple rate is 15.6/1000 live births, compared to an all-Alberta rate of 1.75/1000 live births. Eight of the 16 were born in one postal code region. There were 148 children

born in this region, thus the postal code region rate is 54/1000. Two of the four children who were DE on both first and second samples came from this postal code region. These 16 episodes occurred in the first 145 days of the year whereas in the rest of Alberta the remaining 58 incidents of DE Biotin were spread evenly throughout the year. No other episodes of clustering of abnormal results were found for any of the screened diseases.

There were no statistically significant differences between the mother's age, parity, child's gestation, birth weight, and sex of the 16 children living in Region 1 compared with all other children born in Alberta in 1992.

Tyrosinemia

There were 731 instances where an initial screen reported a high tyrosine level. Of these 731, only 315 GM infants (44.2%) had repeat samples. The average age at repeat was 34.5±21.3 days (range 2 to 236). About 51 percent of second samples were obtained after the child was over 30 days of age. Among the repeats, 14 samples showed a high tyrosine and two were NSQ/US. Of the 14 with a high tyrosine, eight were repeated and found to be normal. Six of these were in infants of low birth weight. The remaining six have no record of a repeat sample. The two with an NSQ/US were not repeated. As well, there were six other infants who for various reasons had a normal initial amino acid screen but who on re-screening by request for other reasons were found to have a TY. There is no record of any of these having a further repeat sample taken.

Congenital Hypothyroidism

There were 59 infants with a initial TSH greater than 25 and who had blood taken after 24 hours of age. In 36 of these infants, serum T4 measures were determined from the original sample; however, all 59 infants had a request to their physician that a repeat sample be obtained within a one month period or sooner if clinically indicated. Repeat thyroid measures were found for only 14 infants, and two of these infants had persisting abnormal results. The second sample was taken at an average age of 13.7±9.6 days (range 2 to 29 days). It is apparent that other repeat samples were obtained, but which were not part of the neonatal screening program, because a total of 11 infants were found who were verified as having congenital hypothyroidism. The author was provided with

this information by the director of the NSP when he was trying to verify the number of children with TSH values greater than 25 who were rescreened. These infants presumably had serum thyroid studies done which would not have been detected with the data retrieval process. All of the infants with congenital hypothyroidism had initial TSH values greater than 50.

NSQ/US Samples

There were 142 NSQ/US samples of which 92 were resampled at a mean of 25.0±28.6 days (range 2 - 270 d; median 21 d). Among resampled cases, one was NSQ, one had an abnormal amino acid and one a decreased Biotin. For the child with the abnormal amino acid, a repeat sample done 22 days later was normal. For the child with the repeat NSQ sample, a second repeat was done 15 days later at age 21 days.

Samples Obtained Before 24 hours of Age

There were 235 instances of 'good' matches where blood was taken before 24 hours of age. Of these 235 children, 109 had repeat samples taken after a mean interval of 9.1±13.8 days (range 1 - 91 days; median = 3 days). Of these 109 repeats, six had a high tyrosine, one a decreased biotinidase, and one NSQ. None of these abnormal results have a record of a further repeat sample. Eighty percent of the 109 had their repeat sample taken by age seven days, but at 30 days ten percent still had not had a repeat sample.

Repeat Samples when the Original was Normal

In 850 instances a repeat value was obtained but the reason for a repeat was not clear as the original screen was normal. Of these, 5.6 percent were abnormal in one respect or other, compared to 2.5 percent among the initial sample group. There were 24 high tyrosine levels, 13 abnormal amino acid patterns, one high phenylalanine level, three decreased biotinidase levels, two high TSH values and five NSQ/US samples. The average age for resampling was 18 days with the mode at three days. Twelve of these children died, 18 percent had a gestational age of 36 weeks or less and 18 percent had a birth weight of 2500 g or less.

CHAPTER 4

Discussion

This thesis has explored the degree to which the Alberta Newborn Screening Program has met its mandate of screening the entire Alberta population of newborn infants in the year 1992. The process employed was not an audit of the screening program as only selected aspects were reviewed. In general, the results describe a program that appears to work well, but with some significant weaknesses, notably in the follow up of abnormal results. Before reviewing the conclusions and because of the potentially negative implications of such results, it is necessary to explore the quality of the data upon which the conclusions are based. For this study, critical issues in data quality are: the source of the screening data, the accuracy of the match, the detection of repeat samples and the accuracy of time estimates when samples were obtained.

Section 1: Data Quality

The Source of the Data

The key source for the screening data was the archived data tapes of the clinical laboratories of the Walter Mackenzie Health Sciences Centre (WMC) of the University of Alberta. These tapes contain all the data for the Newborn Screening Program. Data abstraction, as described in the text, was exhaustive. The author feels comfortable in this assessment because, at one stage of the analysis, it was apparent that data were missing during two specific time periods. Rechecking with the laboratory manager resulted in further data being provided which accounted for virtually all of the "excess" missing infants. Later, in the analytic phase, the author became aware of a second source of data at the Biochemical Genetics Laboratory at the Alberta Children's Hospital in Calgary, but it involved only repeat screen samples and the relevant information regarding specific infants was provided by the laboratory director. The completeness of data collection cannot be assessed directly; however, all the available provincial sources of screen data have been accessed for all disorders except congenital hypothyroidism and, for these infants, if follow up thyroid studies were done in a private laboratory, the resultant data are not available.

The quality and accuracy of the birth data were assumed to be very high, as it is obtained from vital statistics data collected, edited, and provided by the Provincial Government. There is no easy way of verifying its accuracy, but there is no reason to assume significant error.

The Accuracy of the Match

The accuracy of the match cannot be determined precisely but, because of the method of matching it is likely that the accuracy is high. As well, because of the assignment to each record of a level of confidence of a match, the subsequent analysis, which is based mainly on records with strong confidence of a very good match, is conservative in its conclusions. The match was done using deterministic criteria in that matching was based on specific identifiers, such as last name, date and place of birth, and a record either did or did not match that identifier. The hand matching was also mostly deterministic, in that a visual assessment was made of the matching parameters and a decision made. By hand matching, confidence of a "correct" match could be very high. Because there were instances where the match was simply a 'match' drawn from a group of eligible birth records, an assessment of the degree of confidence of the match was made for each record. Initially, a series of levels was considered; however in the end, it was decided to state that a match was a good match (GM), with a very good probability that it was correct, or a poor match (PM), with much less probability that it was correct. The PM infants were not chosen totally at random; rather, they were chosen in a manner that was constrained in that the birth record child had to have a birth date as close as possible but before the birth date of an unmatched screen result, had to be alive on the date the blood was recorded as taken, and had to be from the same community, the next nearest community, or the most likely referral community as the unmatched screen result. Within those constraints, the choice was arbitrary. This partly non deterministic matching makes useful analysis of the PM group debatable and little analysis has been done with these data apart from the estimation of total coverage. Most of the detailed analysis compared the good match group with the not matched group.

The children left over at the end of matching constituted the not matched group.

The non matching appeared to be reasonably well spread across the home regions of the child, with somewhat more non matching occurring among children who were born in Alberta but who lived out of the province.

The Creation of Specific Population Subgroups

In several instances the data were recoded to create new variables identifying specific groups. Most of the time the method of derivation is obvious; for example, infants were grouped by birth weight into <1500 g, 1500 - 2499 g and >2500 g. In the creation of variables describing "Native" status, and "Home Births" the derivation was less obvious. The classification of "Native Status" was done by comparing the place of residence of the infant with a map of Alberta to determine if the residence was on a native reserve or not. For this variable, a child considered Native really is a child who has a high probability of being of American Indian extraction, but, because many natives do not live on reserves, the opposite conclusion cannot be made.

The reason for this classification arises from the knowledge that native individuals have in general poorer health than does the general population. It was considered by the author that this may extend to screening practices as well, and thus an attempt was made to determine the degree to which native infants were well served by the screening program. As well, Streetly (1994) in her studies of screening in England, found that infants of minority groups, especially Black infants, were at greater risk of non screening. The results, to the degree that they can be a valid indicator of native involvement, show that native infants fare no better or worse than does the rest of the population.

For "Home Births", the classification was based on the lack of a hospital identifier in the vital statistics birth data base. These infants were examined specifically because it was felt that they might be at greater risk for non screening and may reflect a risk for a particular life style or attitude by their parents toward medical care. It is possible that some of these births could have occurred en route to a hospital or under other circumstances, and do not necessarily reflect only individuals who made a conscious decision to have a child born at home.

The Detection of Repeat Samples

Repeat samples were detected as repeat because they were labelled as such, or because records with equivalent identifiers were found during the editing, visual matching, or computer matching procedures. While some of the records considered as repeats could have been initial records of children, this is unlikely because the method of matching for repeat screens was always deterministic and almost always verified visually. If it did happen, it would mean that the number of "missed" infants would be less but it would mean also that fewer infants could be accounted for who required a repeat screen. For practical purposes, the search for repeat samples stopped after finding the first repeat, although on occasion several sequential repeats were found for an infant.

Some repeat samples may have been missed because they were not labelled as NEONATAL SCREEN, but for phenylalanine, tyrosine, and biotinidase measurements this is unlikely. These measurements are done only at WMC or at Alberta Children's Hospital in Calgary and both of these sites were searched for data, thus it is most likely that these would have been picked up during data abstraction. For children with high phenylalanine or tyrosine levels, they may not have had a repeat Neonatal Screen but may have had a Pediatric Metabolic Screen instead. This will not change the overall statistics very much. While a high tyrosine may be repeated as a pediatric metabolic screen, most repeats should normally come in as a repeat Neonatal screen; therefore the data regarding repeat tyrosine is likely to be accurate. In the present report, each data file was searched for the term Pediatric Metabolic Screen and, when a child born in 1992 was found, the data were reviewed to determine if this was a repeat screen or a test for some other reason. A repeat thyroid screen may have been missed if a serum sample was submitted for analysis rather than a drop of blood on the screening requisition. There are also other reasons for thyroid repeat samples being missed (see below).

The Assignment of Age of Blood Sampling

An issue not related to the quality of data regarding matching, but of key importance in the assessment of the program's performance is determining the temporal sequence of events for screening. This is important when determining who was screened

too early or too late, and how promptly suspicious results were followed up. A critical aspect of the screening program is that children should have their blood taken after 24 hours of age and before seven days and that if there is a need for a repeat, it should be done either immediately, in the case of inadequate samples or very suspicious results, or when the child is no later than one month of age, in the case of less worrisome results. In this study, age when blood was taken could be determined only indirectly, because birth date was not on the laboratory record and had to be calculated from the laboratory report date and the infant's age at that time. This estimated birth date was used in the match process and when a child was matched, the birth date as recorded on the vital statistics form was used. Thus if the match is wrong, the birth date is likely also to be wrong. For this reason, only GM infants had age calculations made.

A situation where error exists is the determination of which infants had blood taken at less than 24 hours of age. A child was considered to have blood taken at less than 24 hours of age if the birth date and the date of blood taking were the same. In all of these instances, the child would be less than 24 hours old; however, if a child were born at, say 20:00 on July 1, 1992 and had blood taken at 16:00 on July 2, 1992, that child would be considered to have had blood taken at greater than 24 hours of age when in reality it was only 18 hours old. While the screen record often stated the time when blood was taken, the birth record did not state the time of birth thus exact calculation was impossible. The estimate of 235 infants whose blood was obtained at less than 24 hours almost certainly underestimates the true situation; however, it is possible to consider with certainty these infants as all having blood taken at less than 24 hours.

The Value of this Research

This research is of particular value because data from a Neonatal Screening Program have been matched with corresponding birth data. This has resulted in a more precise estimate of screening coverage as well as determination of some of the demographic characteristics of infants who were not screened. While similar matching has been reported at other sites, it is uncommon. To the author's knowledge this is the most extensive assessment of screening coverage done to date. No other study has reported the

characteristics of unscreened infants as completely. As well, these data, because they describe the Alberta situation, provide a solid foundation upon which to compare the functioning of the Alberta Newborn Screening Program at other time periods, particularly after 1994 when the provision of health care has changed dramatically and when there is concern by the director of the program that the level of coverage is less than it was in 1992. Finally, this document also points out the likely "problem areas" in data collection and analysis and also the groups at high risk for an "irregular screening event", i.e. non screening, or non-follow-up of unsatisfactory or abnormal results.

Section 2: The Results

Coverage

The first objective of this project was "*To determine the proportion of liveborn infants who were not screened within seven days of birth for the diseases phenylketonuria, congenital hypothyroidism, tyrosinemia, and biotinidase deficiency.*" While sufficiently accurate for administrative purposes, the estimates of coverage provided by most programs are crude. They usually just compare total unique screens and total births for the same area and time. This may introduce a false sense of security in that, if the count of unique screens is incorrectly high, coverage is overestimated. Using the above technique, Dr. F. Bamforth, Director of the Alberta Newborn Screening Program, estimated that in 1992, 98.3 percent of infants were screened. This study extended the usual estimate of coverage by linking screen records with birth data and determining actual coverage and showed that 98.0 percent of children were screened at least once; this is quite a good agreement. However, the relatively small value of two percent missed must be translated into the absolutely large number of 839 infants who were not screened at all in 1992. The two percent missed does not include a further 531 were first screened only after seven days of age, nor does it include 47 infants who had inadequate or unsatisfactory (NSQ/US) samples and who were not rescreened. While the Newborn Screening Program analyses NSQ/US samples, they do not report them if they are normal; instead they request a repeat sample. For this reason, these 47 are not included with the 839 infants with no apparent screening, but the argument could be made that they should be.

Total coverage is important. Recently, in Alberta a child with PKU was missed in the newborn period and not detected until the child presented with developmental delay at nearly a year of age. PKU is a disorder with an incidence of 1:15000. Congenital hypothyroidism is nearly four times as common, with an incidence of 1:4000 (Allen, Hendricks, Sieger et al, 1988). Therefore, on average, about once every four years, a child with CH will be missed in Alberta just because he or she was not screened. The costs to society for the lifetime care of that child, and the loss of contributions by that

child to society, may well pay for the cost of the Newborn Screening Program for that 4-year period, to say nothing of the personal and emotional costs to that child and his or her family. PKU will also be missed, although only about once every 15 years, but with equally tragic consequences.

For these reasons, efforts are required to ensure 100% coverage. While many jurisdictions have policies that children should have screening samples obtained before discharge, errors will occur and some sort of fail safe mechanism is required. Thus, for example, the Newborn Screening Program in British Columbia employs a secondary surveillance system to ensure newborn screening is done[1]. This is done through linkage with the hospitals such that a child's record is not closed until the newborn screen report is attached. If this report is not forthcoming within a month, the hospital lets the screening laboratory know and then the physician is contacted to get the child screened. This will not pick up untested infants born out of a hospital, but it will pick up infants that were not screened in a hospital. Such a program depends on the Medical Records' department of a hospital to report non screening on an obsessive and regular basis but it would help to ensure more complete coverage.

In Britain there have been attempts to inform daily the screening laboratory of births occurring in the past 24 hours. The laboratory then expects to obtain a screen sample from the infant and if it is not forthcoming, then an effort will be made to find the infant and test it. In 1995 New Brunswick instituted a similar program. All hospitals in the province fax the newborn screening program laboratory a daily accounting of all infants born in the previous 24 hours. This information is entered into a computer data base and is linked to the screen samples when they arrive. If a sample is not forthcoming, the NSP contacts the child's physician. If this fails, the director of the NSP attempts to find the child[2]. The New Brunswick system should be able to be adapted to work in Alberta which currently has no surveillance program. Modern communications should permit even the most remote parts of Alberta to inform the laboratory of births when they

---

1. H. Vallance, Director, Newborn Screening Program, British Columbia, personal communication.
2. S. Sanderson, Director, Newborn Screening Program, New Brunswick, personal communication

occur. In Alberta, with the introduction of a lifetime Personal Health Number for each resident such a system is theoretically feasible if the infant were assigned that number at birth and if the laboratory were informed of the birth.

Comment

The key conclusion to come out of the above discussion is that there is a need for a specific screening program protocol that will ensure 100 percent coverage of newborns and rapid follow up of all samples needing repeat. Exactly how this would be designed, implemented, and evaluated requires further thought.

Who Got Missed

A major strength of this thesis is the description of the characteristics of infants who were not screened. This has allowed it to attempt to meet the second research objective: *"Using the available demographic and biological data, to define the characteristics of infants who were not screened, or who could not be identified with confidence as being screened."* Unfortunately, only those data that were recorded on the screen or birth records were available for analysis. Within those limitations, the results here are similar to those of Streetly et al (1994) who suggested that it was infants in special care units, infants who were born in one area but normally lived in another, and infants of high risk ethnic groups that were most likely to be missed. The current data verify that low birth weight infants are at significant risk of being missed, and extend these findings to show that risk of being missed is also associated with single parenthood, home birth, and death in the first week of life. In this study, neither mobility nor native status was particularly associated with being missed.

Death in the first week of life is the strongest predictor of non screening, but, since death occurred mainly in the first few days of life this finding is not surprising and is not of much practical use. At least one author[1] feels that all infants should be included in the assessment of screening coverage and in this study all infants were included. Because of this, 140 of the 167 deaths by age seven days were in the unmatched group. Many of these infants also were of very low birth weight and many also likely very sick. Logistic

---

1. Allan Guttmacher, personal communication

regression showed that retaining the variable "death in week 1" in the analysis did not eliminate the conclusion that low birth weight infants who survived were among those at greatest risk of not being screened.

The next most important predictor was being born out of a hospital, presumably a home birth. Of the 226 infants who were born out of a hospital, 58 were not screened. Again, while this points out a high risk group, it is not a particularly big one, accounting for only seven percent of the NM group. There was no basis in the literature to consider that infants born out of a hospital would be at any greater risk for non screening but the results of this study suggest that they are nearly 20 times more likely to be not-screened than are children born in a large hospital. Nearly 25 percent of these infants were not screened, and of those that were screened, many had the blood taken late, when the infant was more than seven days of age. Infants born out of a hospital were on average heavier than infants born in a hospital and their mothers were slightly older and more likely to be multiparous. Why these children are more likely to not be screened is not clear but may reflect a lifestyle decision on the part of the parent.

Numerically more important were low birth weight infants. Among those who did survive to age one, it was this group of infants who was at greatest risk of not being screened. Finally, infants of mothers who were divorced, or widowed, or separated, or whose marital status was unknown were also at high risk of not being screened. Infants of single mothers were at lesser risk and infants of mothers who were married were at least risk.

Some of these results seem opposite to what one would expect. This is particularly so for the finding that low birth weight infants are at particularly high risk. These infants normally have care provided by specialists who should have a high level of sensitivity to metabolic disease; however, for some reason these infants are quite often missed. It could be that a "simple thing" like a neonatal screen is forgotten in the anxious early days of life of a sick newborn, or it could be that there is no fixed policy that a screen be done before discharge, or it could be that the child is transferred to another institution and the assumption is made by the receiving institution that the screen was

already done. The problem of small babies being missed clearly needs to be addressed in an education program. On the other hand, if a very low birth weight infant was screened and had an abnormal result, the odds of having a repeat sample were 2.9 times those of a normal weight infant needing a repeat screen (Table 3-21). For clarity, the odds shown here are inverted from those seen in the table. This seems to suggest that if a screen is thought about, any necessary follow-up is appropriate.

Another unexpected result is the strong association of being unmatched with mothers whose marital status was divorced, separated, widowed, or unknown, but a less strong, but still significant, association with single mothers. Formerly married mothers are 10 times as likely as married mothers to be unmatched, but single mothers are only 2.2 times as likely to be unmatched. If the odds ratios were similar one might assume that not matching is associated with lower socioeconomic status, but they are not. It is difficult to reconcile this observation. It could reflect an error in matching due to conflicting surnames on the various records. The birth record contained the surname of the child (which commonly is the same as the father) and mother, but not the father. The screen record only had one surname listed. Matching was usually based on the child's surname but if that failed, then the mother's surname was used. If the screen record used the name of the father or mother, a match would normally eventually be made. If the screen record used another name, a match would not occur. The question remains though: "Why another name?"

While other potential determinants of non screening were examined, including maternal age, gender, Native status, movement between communities, and health administrative region, these factors were not generally predictive of non screening. Simple categorical analysis had suggested these as possible factors for non screening but logistic regression analysis removed them as predictors.

Comment

The logistic equation was quite efficient at defining the characteristics of who was at risk. Of the 802 unmatched infants for which all the data in Table 3-13 were available, 630 had one or more risk factors as determined by logistic regression. While this may

suggest that those infants at risk of being missed can be identified, it would be a false conclusion as 172 (21.4%) had no risk factor and yet were missed. Regardless of the presence or absence of risk factors, most children are screened and while some children are at greater risk of being not screened, one can be certain a child has been screened only if that child can be linked to a screen result.

Timing of Initial Samples

Ideally, samples should be obtained between one and seven days of age and, for most infants this was the situation in Alberta, with less than 1.9 percent of infants having blood obtained outside this desired range. Infants having samples taken too early may result in a false negative result for PKU and a false positive result for TSH. The data support this in part. None of the 235 infants sampled at less than 24 hours or at one day of age had a high PKU on the initial sample or on resampling, when done. On the other hand, 11 of the 235 (4.26%) infants sampled at less than 24 hours had a high TSH, whereas only 11 of 1527 (0.72%) infants who were sampled at one day of age and 46 of 38781 (0.12%) infants sampled after one day of age had high TSH values. These data show that TSH values are higher earlier in life and may reflect false positives. The high incidence of abnormal results in the less than 24 hour group points to a problem that will increase as early discharge becomes more common. There will be a greater need for re sampling, and therefore a greater need to follow-up abnormal results and to identify screening samples as repeats. Because of the slightly higher proportion of a high TSH on day one, and since TSH will usually be within the normal range by day one, the data also suggest that there may have been some infants classified as one day old infants who were actually sampled when less than 24 hours of age.

Samples taken after seven days of age may result in affected children not receiving treatment as soon as possible. There were 38 children who had the initial sample taken after seven days and had results requiring repeat. Of these, 17 were repeated and all were normal, but nothing is known about the remaining 21 children. The average age of repeat was 30 days (median 20 days) and the average interval between first and second samples was 19 days. These values are not significantly different from the average age of repeat

for all children who obtained required repeat samples.

Repeat Screens

The third objective of this project was *"To determine the degree to which infants who when screened and had results requiring follow up, were followed up appropriately and expeditiously and a final disposition made."* The assumption made in this report is that if a child had an abnormal screen result or an unsatisfactory sample, or a sample taken before 24 hours, then there was an absolute requirement to re sample. The speed and completeness of this process reflect the effectiveness of the program as it is these children who are likely to be at greatest risk for metabolic disease. There were few features characterizing infants who needed a repeat screen. The logistic regression results suggest that it was the small or sick infant who was most likely to need a repeat. Native infants also were twice as likely as non-native infants to require a repeat. The reason for this appears to be an increased frequency of NSQ/US results.

Unfortunately, follow up of repeat screens is an area where the Newborn Screening Program did not function as effectively or as efficiently as it should have. The most important finding of this study with respect to repeat screens is that only half of the infants who needed a repeat actually had it done. For infants who were "good" matches, there were 1375 cases where a repeat sample was requested, but in only 662 infants (48%) was a repeat sample verified as taken. Fortunately, among infants with high phenylalanine levels, all had a repeat obtained, but only 10 of 11 infants with a TSH value of >50 had a repeat done. Finally, among those infants who did have the required rescreening done, the time frame for repeat sampling often was greater than 30 days, and the benefits of early intervention, if necessary, were beginning to be lost. The logistic regression analysis shows that very low birth weight infants are significantly more likely to have a needed repeat done than are normal weight infants. This may reflect the longer stay in hospital for low birth weight infants, and thus easier access.

This figure of 50 percent seems low, but there are no ready figures for comparison as to what proportion of re-sampling constitutes good practice. In order to develop some comparative standards the author obtained some data from various screening programs in

Canada, the United States, Europe, and England to determine their rates of coverage and of re-sampling response. The general impression from reviewing these programs is that where there is a clear, time oriented, protocol to ensure complete follow up of repeat requests, compliance for rescreening is high, but where there is no specific protocol to manage repeat requests, coverage is poor.

For Ontario, the program is voluntary. The Ontario Ministry of Health provides a screening guide but screening policy is at the provider level. While all hospitals attempt to ensure that all infants are screened, the level of initial coverage and of repeat screens is difficult to determine. For abnormal results, "The laboratory staff inform the regional consultant who then notifies the family physician for a follow up...The laboratory staff **ensure** that a repeat sample reaches them. "[1] British Columbia estimates that a second sample is received in approximately 95 percent of cases. The missing five percent are almost totally only those that were less than 24 hours and did not get rescreened[2]. Quebec estimated that in 1992, 98.5 percent of infants were sampled (Screens/Births). Of these, 384 of 95812 samples were inadequate and 277 were repeated (72%). There were 166 abnormal results, and all were re-sampled[3]. In 1995 Manitoba had 16518 first screens. This is more than the total number of births. Of these, there were 814 repeat requests and all except for 81 were received. Manitoba has a specific protocol for abnormal results which involves verification of results on the same sample, notification of the responsible physician by phone or by mail depending on the urgency of the situation, and pediatric consultation as needed. If no follow up is received within 2 weeks of the original request, the public health nurse is notified who contacts the family and arranges for collection. Public health was involved in 184 of the 814 repeat screens mentioned above[4].

---

1. Dr. HP Demshar, Director, Lab. Services Branch, Ontario Health. Personal communication. The bold is his emphasis.

2. H. Vallance. Director, Newborn Screening Program, British Columbia. Personal communication.

3. A. Grenier, Quebec Neonatal Screening Program. Personal communication

4. I. Wilkinson, Head, Clinical Chemistry and Metabolic Diseases, Cadham Provincial Laboratory, Manitoba. Personal communication.

In Vermont the proportion of repeat screens is nearly 100 percent, largely because of aggressive follow-up of abnormal results[1]. In California, genetic centre nurses follow all repeat requests until the second test is obtained. Usually this is done by phone, and usually it is done through physician contact. Parents are not normally contacted, but in extreme cases, they may be, as may be Child Protective Services[2]. In Oregon state, the child's physician must ensure that a screening sample is drawn, and the Pacific Northwest Regional Newborn Screening Program provides an education program about screening to hospitals, birth centres, nurses, lay midwives, and physicians. It also has a monthly Screening Practice Profile which lists the hospital's problems (if any) with specimen collection and handling.

These programs seem to work. In 1994, about 99.7 percent of infants born in Oregon had an initial specimen drawn, as estimated from a count of births vs unique screens. About four to five percent required a repeat screen because of abnormal results or unsatisfactory specimens and only 10 infants were "lost to follow up"[3].

In Switzerland, because of confidentiality laws, screening coverage cannot be assessed by comparing birth registry data directly with screening data, however, total screens almost always exceed total births, even after accounting for known repeats and for out of country residents, thus coverage is estimated to be 100 percent. Repeat sampling is managed through the family doctor and compliance is very close to 100 percent[4]. Switzerland has a computer system which registers the recalls requested and which produces a daily list of outstanding samples. If there is a delay of >7 days, a reminder is sent. This program also supplies the screening form and pre-stamped and pre-addressed envelopes. In other words, they do their best to make it easy.

The United Kingdom has several different screening programs, based on a regional system. In the Trent Neonatal Screening Service, initial coverage varied with the district,

---

1. A. Guttmacher, Director Vermont Neonatal Screening Program, Personal communication.

2. In California, the loss to follow up for sickle cell disease is about 2 percent; for PKU the loss rate for initial positives is about 0.4 percent; hypothyroidism is about 0.3 percent (F. Lorey, California Neonatal Screening Program, email communication).

3. J. Tuerck, Pacific Northwest Regional Newborn Screening Program, personal communication.

4. T. Torresani, Personal communication.

but for those infants where a re-screen was requested in 1995 because of an abnormal result, (not an inadequate sample), all complied[1].

Poor compliance for repeat screening is not unique to Alberta. Ms J Tuerck of the Oregon Screening Program is a member of the Maternal and Child Health Select Panel on Newborn Screening Systems which has evaluated 14 states in the US and she comments that:

> *"follow up problems seem to be common to all the states. I am sure it is due to the relative lack of attention over the past 30 years. Follow up which works best seem to be those who have a designated person(s) at the program manager level, use medical consultants for the significant abnormals, have written protocols which are followed and evaluated on a regular basis, and who have close effective communication with other parts of the program (lab and practitioners) and with other state or provincial programs."*[2]

Several explanations may account for the poor performance in Alberta. There is no specific protocol for follow up of repeat requests, repeat screen samples are poorly identified, and repeat samples are sometimes done in private laboratories. A key reason for the poor performance in Alberta is the lack of a clear, time oriented, protocol for dealing with follow-up requests and the lack of personnel dedicated to this specific task. While abnormal results are logged and requests for a repeat sample are also logged, there is no time oriented mechanism to ensure follow up. While it is obvious from the results of this study that serious metabolic disease, such as PKU and CH are being detected, and that follow up of some sort is taking place, it is not highly structured and the possibility of an affected child being missed because of failure to repeat a sample is real.

A second factor contributing to the apparent poor follow up is that repeat screenings are not necessarily labelled as a repeat and thus the laboratory may not identify the sample as a repeat. This was evident to the author during the data abstraction process, when on occasion, a note would be appended to a report stating that this was a repeat screen. On many occasions, however, no such note would be present, indicating that the

---

1. R Pollitt, personal communication.
2. J. Tuerck, personal communication

lab had not recognized this as a repeat screen. This explanation is not particularly applicable in the present instance, because during data abstraction, repeats were identified, even on occasions when the surname had changed (say from the mother's surname to the father's). It does point out, however, the difficulty the screening program has in identifying repeat screen samples.

A third reason for poor follow-up may be a lack of knowledge on the part of the physician as to what to do with the report, or an unsatisfactory attitude towards its importance. Some physicians, upon reading a message in the screening report "*SUGGEST REPEAT SCREEN IN ABOUT 1 MONTH OR SOONER IF CLINICALLY INDICATED*", may assume that if they think it is not clinically indicated, the sample need not be repeated. This is a wrong interpretation. These diseases can not be eliminated with certainty on the basis of clinical impression only; non-disease requires laboratory verification.

A final reason for apparent poor follow up is that, while most repeat samples were done only at WMC, or at the Alberta Children's Hospital in Calgary, some thyroid samples could have been repeated in any private lab and no record of a repeat would be obtained. This likely accounts for the fact that repeat results for only 24.1 percent of infants with high TSH values could be obtained. In telephone conversation with the laboratory managers of the private laboratories in Edmonton and Calgary it is the author's understanding that if the sample is known to be related to the newborn screening program, it is sent to the WMC in Edmonton for analysis. If a serum sample is taken for analysis, rather than a drop on a filter paper, the laboratory often will do the analysis itself. Given that congenital hypothyroidism is the most common of the disorders screened for, and that treatment is most successful if it is started early and monitored by an endocrinologist, this very poor follow-up is extremely worrisome and underlines the need for an individual dedicated to ensuring that children with abnormal results are followed up expeditiously and completely.

Section 3: Specific Results

While most of this research has dealt with the NSP as a whole, some comments relating to specific results are in order, as they exemplify particular aspects and problems of the screening program.

Phenylketonuria

Phenylketonuria was the "original" disease for newborn screening programs and, except for the example mentioned below all infants with high phenylalanine levels were followed promptly and a diagnosis of PKU made in two infants, who were started on therapy well before one month of age. One child was detected on initial screen to have a high phenylalanine level and a repeat was requested. Two repeat samples were obtained and each time the result was abnormal, but no contact was made with a specialist in metabolic disease or with a metabolic dietitian. The child was then lost to follow-up and no one was aware of his potential problem until the author contacted the laboratory to inquire of the disposition of this child. Review of the case in 1996 shows that the child probably had a benign variant of hyperphenylalaninemia but should have been followed with periodic measures of phenylalanine. This case illustrates the need for a clear protocol of follow up until resolution and also illustrates the need for expert evaluation of all infants with persistently abnormal results.

Congenital Hypothyroidism

With an incidence of approximately 1:3500 live births, congenital hypothyroidism is the most common of the metabolic diseases screened for. It is also a disease where speed of diagnosis and early therapy are essential for optimum outcome. It also is the disease where false positive results are most likely to occur if the sample is obtained before 24 hours of age. For all these reasons, it is essential that a strict protocol for follow up of suspect results be employed and yet it was this disease where follow up was the most problematic, largely because a repeat sample may be analysed at a private laboratory and the result not forwarded to the screening program. A TSH value of over 50 is an indication of a very strong likelihood of congenital hypothyroidism. For the eleven infants whose TSH level was over 50, 10 were retested and all found to have CH.

For one infant, with a TSH of 54, there is no record of a repeat screen. For the 47 infants whose TSH was between 25 and 50, the possibility of hypothyroidism is less, but results still require verification; however, a repeat sample was found for only 12 of these infants. This poor level of follow up is not acceptable and again underscores the need for a mechanism to ensure follow up.

NSQ/US Results

There were 142 instances where the sample was inadequate or unsatisfactory. Of these, only 46.4 percent were repeated, however, among the repeats there were several abnormal results which required verification. Where data are available, all of these eventually resulted in normal values being obtained but it does point out the need to follow up children with NSQ/US samples.

Most of the NSQ/US results came from the larger centres, where most samples were obtained. Relatively more of the NSQ/US results came from centres where the frequency of sampling was once a week or less, suggesting that an educational program may be needed to ensure appropriate sampling technique. There are also regional differences in the frequency of NSQ/US samples, with one region having over four percent of all samples being of poor quality. The next closest region had less than two percent of samples being NSQ/US.

Tyrosine

A raised tyrosine level was the most common reason for an abnormal screening result, occurring in 714 infants for whom age at screening was known; this is 1.8 percent of all those screened after 24 hours of age. There are several rare syndromes associated with high tyrosine levels that are partially or totally amenable to therapy and that have significant consequences if not diagnosed; however, the most common cause of a raised tyrosine is transient tyrosinemia due to metabolic immaturity. This is often seen in premature infants and in this study, 163 (22.8%) of the GM infants with high tyrosine had birth weight of less than 2500 g, whereas only 5.5 percent of all GM infants weighed less than 2500 g. The presence of raised tyrosine levels in two percent of neonates may result in a sense of false security and a less than optimal obsession with obtaining a repeat

specimen. When a high tyrosine level was found, the laboratory usually did a succinylacetone measurement on the same sample. A normal result meant there was no Tyrosinemia Type I disease, but does not exclude Tyrosinemia Type II which is rare (1:50,000) but amenable to therapy, nor does it exclude transient neonatal tyrosinemia, a disorder with possible long term consequences, but likely for only a subset of infants (Scriver, Beaudet, Sly Valle, 1989). Therefore, if there is a high tyrosine result, there is a requirement for a repeat sample.

Biotinidase Deficiency

The most striking finding in the study with respect to Biotinidase deficiency was the very high incidence of "DE" results arising in Taber. It was noted by the NSP in 1992 and an investigation was done but no specific answer was found. If the sample is exposed to a bleaching agent, it may give a false positive. All the samples came from one hospital thus it is possible that they may have been contaminated. This was felt to be most likely, as after a review of procedures and ensuring that samples were properly stored, there were no more abnormal results from that site.

Repeat Samples That Were Not Indicated

There were 884 instances where a repeat sample was obtained but where the initial sample was normal. Among these 884, 33 had abnormal or NSQ/US results. Why these screens were obtained is not clear. The repeat screens were obtained at an average age of 20.45 days (median 7) and 17.5 days (median 5) between the first and second samples. Cross-tabulation analysis of the data suggests that these samples were taken from children who were at risk or who were born in one region but who normally resided in another. Nearly 28 percent of GM children who had an "unnecessary" repeat had a birth weight of <2500 g but only five percent of GM children with no repeat had a birth weight of < 2500 g ($\chi^2 = 745$, $p < 0.05$). It seems likely that many of these children were graduates of an newborn ICU and received a second screen prior to discharge. Mobility *per se* did not appear to account for an unnecessary repeat; however, being born in a region different from the home region was significantly associated with extra repeats ($\chi^2 = 24.3$, $p < 0.05$). It is possible that some of these "unnecessary" results were in fact repeats that were

needed but were not recognized as such due to linkage problems. While possible, this is not felt to be a likely reason as all of these cases were visually examined to determine if they could be linked to other records.

Section 4: Conclusions

This research has combined birth registration and screening data to determine the level of coverage of a newborn screening program and to describe features characteristic of non-screened infants. No other study, as far as the author can determine, has compared the screened population to birth registry data on as huge a scale; most reviews have dismissed the idea as unfeasible. As well, this study has determined the characteristics of missing infants so that educational packages can be made to most effectively reach those at greatest risk of being missed.

This study has provided an opportunity to examine the Alberta Newborn Screening Program and to make suggestions where appropriate that could improve the program. This research is needed also to address the effectiveness of the current Newborn Screening Program in Alberta, especially with respect to follow-up procedures. With early discharge, there will be a greater number of infants requiring follow-up and some mechanism will have to be developed to ensure that it is complete, definitive, and timely. The data from this study will help in the planning such a system.

This project has examined the degree of coverage of the Alberta Newborn Screening Program for the year 1992. Nearly 98% of newborn infants were screened for metabolic disease, leaving 839 infants unscreened. Low birth weight infants, infants born out of hospital, and infants born of unmarried mothers were most commonly missed. Among infants requiring a repeat screen for any reason, nearly 50 percent did not have the required repeat done. The reason for this is most likely due to a lack of a specific protocol detailing the actions necessary when there is an abnormal screening result.

Some Suggestions Regarding the Alberta Newborn Screening Program

The following have arisen out of the findings of this research, as well as the literature describing screening programs elsewhere in the world. They are not exhaustive, but they do represent areas that the author feels should be addressed promptly.

1. Efforts need to be made to ensure that all infants are screened at between one and seven days of age. Particular effort should be directed to infants who are in a special care nursery and to infants who are born at home. The former could be done by improving educational programs to those working in special care nurseries and the latter by ensuring that those who deliver infants out of hospital are informed as to the need for early screening.
2. A clear protocol must be developed that details the procedures for the initial process of screening and the subsequent process of follow up of suspicious or unsatisfactory results. This protocol should include techniques for continuing evaluation of program performance, particularly initial coverage and follow up.
3. There should be an individual whose principal task is to ensure that the program runs effectively and efficiently. On average, every working day there will be 9 children who will either be missed, be late in screening, be too early in screening or have an abnormal or uninterpretable result.
4. All infants should be provided with a unique health care identifier at birth. If assignment were done centrally, the same process could inform the newborn screening program where and when an infant was born, what the health care number was, and the NSP could then anticipate receipt of the screen. This would allow follow up if receipt of a sample was slow.
5. There needs to be an educational program to physicians, midwives, public health personnel, and new parents regarding the need for newborn sampling. In a similar manner, there should be a program of feedback to physicians and hospitals with respect to the functioning of the newborn screening program. There should be inservice regarding the technique of obtaining good samples.

REFERENCES

Allen DB, Hendricks A, Sieger J, et al, 1988: **Screening programs for congenital hypothyroidism**. Amer J Dis Child 142:232-236.

Brown A, Lyon I, 1987: **Multiple births and hypothyroid screening**. In: Therrel BL ed: *Advances in neonatal screening*. Elsevier Science Publishers, pp 63-64.

Committee on Genetics, 1992: **Issues in Newborn Screening**. Pediatrics, 89(2):345-349.

Connelly J, 1987: **Australian experience in neonatal thyroid screening**, 1977-1985. In: Therrel BL ed: *Advances in neonatal screening*. Elsevier Science Publishers, pp 31-34.

Dhondt JL, Farriaux JP, et al, 1991: **Economic evaluation of cost-benefit ratio of neonatal screening procedure for phenylketonuria and hypothyroidism**. J Inher Metab Dis, 14:633-639.

Dobbins RH, 1987: **Evaluation of neonatal TSH data for 6 Age categories**. In: Therrel BL ed: *Advances in neonatal screening*. Elsevier Science Publishers, pp 575-576.

Dussault JH, Coulombe P, Laberge C (et al), Letarte J, Guyda H, Khoury K, 1975: **Preliminary report on a mass screening program for neonatal hypothyroidism**. J Pediatr, 86:670-674.

Ferraira, P (1985): **Current Canadian Newborn Screening Practices**. Pediatric Medical Quarterly: pp 111-122.

Fisher DA, 1987: **Effectiveness of newborn screening programs for congenital hypothyroidism: prevalence of missed cases.** Ped. Clin. N Am, 34(4):881-90.

Foley BL, Fisher DA, Shapiro LJ, Cunningham GC, 1987: **Early neonatal discharge and its effects on thyroid screening results in southern California**. In: Therrel BL ed: *Advances in neonatal screening*. Elsevier Science Publishers, pp 55-66.

Fyro K, Bodegard G, 1987: **Four-year follow-up of psychological reactions to false positive screening tests for congenital hypothyroidism**. Acta Paediatr Scand. 76:107-14.

Griffith PD, Morris J, Assheton J, Green A, 1987: **An on-line computerised system for neonatal screening.** In: Therrel BL ed: *Advances in neonatal screening.* Elsevier Science Publishers, pp 513-516

Gruters A, Delange F, Giovannelli G, et al, 1994: **Guidelines for Neonatal Screening Programs for Congenital Hypothyroidism.** ESPE Working Group on Congenital Hypothyroidism. Horm Res. 41:1-2.

Guthrie R, Susi A, 1963: **A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants**. Pediatrics, 32:338-43.

Hisashige A, 1994: **Health economic analysis of the neonatal screening program in Japan.** Int. J. Tech Assess in Health Care, 10:(3)382-391.

Holtzman NA, McCabe ERB, Cunningham GC, et al, 1981: **Screening for phenylketonuria**. New Eng J Med. 304:1300.

Holtzman C, Stazyk W, Cordero J, Hannon H, 1986: **Descriptive Epidemiology of Missed Cases of Phenylketonuria and Congenital Hypothyroidism**. Pediatrics, 78:553-558.

Irie M, Nakajima H, Inomata H, Noruse H, Suwa S, Takasugi N, 1987: **Screening of neonatal hypothyroidism in Japan**. In: Therrel BL ed: *Advances in neonatal screening.* Elsevier Science Publishers, pp 41-47.

Italian Committee of Experts to Neonatal Mass Screening, 1987: **Neonatal mass screening for metabolic disorders in Italy: Technical report of Italian Society of Pediatrics**. In: Therrel BL ed: *Advances in neonatal screening.* Elsevier Science Publishers, pp 473-474.

LaFranchi S, Dussault JH, Fisher DA et al, 1993: **Newborn screening for congenital hypothyroidism: Recommended Guidelines.** AAP Section on Endocrinology and Committee on Genetics, and American Thyroid Association Committee on Public Health. Pediatrics, 91(6):1203-1209.

Last, JM ed, 1985: **A Dictionary of Epidemiology**. Oxford University Press, New York, pp 152.

Leger J, 1990: **Screening for congenital hypothyroidism in France. Misdiagnosed cases: collagorative study of screening centres in France.** European J. Ped, 149(9):605-607.

Listernick R, Frisone L, Silverman BL, 1992: **Delayed diagnosis of infants with abnormal neonatal screens.** JAMA 267(8):1095-1099.

MacMillan DR, Mabry CC, 1995: **Differences in screening and confirmatory thyroid profiles in congenital ypothyroid subtypes: implications for missed cases.** J Perinatology 15(2):126-130.

Mamunes P, Prince PE, Thornton NH et al, 1976: **Intellectual deficits after transient tyrosinemia in the term neonate.** Pediatrics 57;5:675-80.

Newborn Screening Committee, 1995: The Council of Regional Networks for Genetic Services (CORN), **National Newborn Screening Report - 1992**, CORN, Atlanta.

Pass KA, Schedlbauer LM, MacCubbin PA, Glebatis DM, 1991: **Comparison of newborn screening records and birth certificates to estimate bias in newborn HIV surveys.** Amer J Pub Health,81, Supplement 22-24.

Pharoah POD, Madden MP, 1992: **Audit of screening for congenital hypothyroidism.** Arch Dis Child, 67:1073-6.

Rothenberg MB, Sills EM, 1968: **Iatrogenesis: the PKU anxiety syndrome.** J. Am. Acad. Child Psychiatry, 7:689-92.

Saslow JG, Post EM, Southard CA, 1996: **Thyroid screening for early discharged infants.** Peditrics, 98;1:41-44.

Schoenberg D, Klett M, 1987: **Screening for congenital hypothyroidism in the Federal Republic of Germany. Past, Present, and Future.** In: Therrel BL ed: *Advances in neonatal screening.* Elsevier Science Publishers, pp 25-29.

Scriver CR, Beaudet AL, Sly WS, Valle D; consulting editors, JB. Stanbury, JB Wyngaarden, DS Fredrickson, 1989: **The Metabolic Basis of Inherited Disease**, 6th ed. McGraw-Hill; pp 556.

Sinai LN, Kim SC, Casey R, Pinto-Martin JA, 1995: **Phenylketonuria screening: effect of early newborn discharge.** Pediatrics 96(4 Pt 1):605-608.

Smith I, Cook B, Beasley M, 1991: **Review of neonatal screening programme for phenylketonuria.** BMJ, 303:333-335

Sorenson JR, Levy HL, Mangione TW, et al, 1984: **Parental response to repeat testing with "false positive" results in a newborn screening program.** Pediatrics, 73:183-187.

Streetly A, Grant C, (et al) Bickler G, Eldridge P, Bird S, Griffiths W, 1994: **Variation in coverage by ethnic group of neonatal (Guthrie) screening programme in south London.** BMJ 309:372-374.

Verkerk PH, Vaandrager GJ, Sengers RC, 1990: **Fifteen years of national screening for phenylketonuria in the Netherlands;** 4th Report of the National Commission for Management of Phenylketonuria [Dutch}-abstract. Nederlands Tijdschrift voor Geneeskunde 134(52):2533-2536.

Widhalm K, Wintersperger U, Bischof S, Brix R, 1995: **Screening for Biotinidase deficiency in Austria.** Screening 4:71-78.

Willi SM, Moshang T, 1991: **Diagnostic dilemmas. Results of screening tests for congenital hypothyroidism.** Ped Clin N Amer, 38(3):555-566.

Wilson JMG, Jungner G, 1968: **Principles and Practice of Screening for Disease.** WHO, Geneva.

Wolf B, Grier RE, Allen RJ, Goodman SI, Cein CL, 1983. **Biotinidase deficiency: an enzymatic defect in late-onset multiple carboxylase deficiency.** Clin Chim Acta, 131:273.

APPENDIX 1: DATA ABSTRACTION

Abstracting the Data from the Laboratory File

Data containing the screening results were stored on archival tapes in the Department of Clinical Laboratories of the Walter Mackenzie Center. Ten laboratory data files, providing all the clinical laboratory data for 1992 and for the first 2 months of 1993, were provided to the researcher by the laboratory manager, Ms. Kathy Mazurek. The files were copied to a personal computer and the relevant screen data were then abstracted. It was necessary to use a custom written computer data abstraction program to read the "lab file" and search out and write to a separate "screen file" only those data that related to a newborn screen.

Cleaning and Verifying the Screen File

The resultant "screen" file required several stages of editing before being matched with the birth data. The first stage was to edit the screen file by examining it and determining obvious areas where the lab file had to be visually examined and necessary corrections made to the abstracted record. In this stage, it was obvious errors or omissions in the screen file that dictated where to look in the laboratory file for verification or correction. For the remaining stages, it was a search through the laboratory file that dictated changes to the screen file. These stages consisted of several searches of the lab file to:

1. Detect any missing cases;
2. Obtain any added comments that the lab had written (such as a suggestion to repeat the sample, or that the sample was obtained before 24 hours.);
3. Detect twins;
4. Look for known name changes; and
5. Detect any home births.

These searches were done by using a text editor and searching the lab file for specific phrases that were commonly appended to the end of the laboratory report and reflected the comments of the laboratory about the result. They often were requests for a further newborn screening sample to be obtained within a given time. On occasion, they were

messages regarding the sample quality, or the fact that the child's name had changed, or other issues.

The search phrases included:

{TWINS}|{#?,\}|{"A"}|{"B"}|{"C"}|{'A'}|{'B'}|{'C'}|{HOME*BIRTH}|{MID*WIFE}

This means, search for the term 'TWINS' *or* any phrase consisting of a number, '#', followed by a character, '?', a comma and then a space, '\ ', *or* any 'A' (*or* 'B' *or* 'C') in quotation marks *or* the words 'HOME' followed on the same line by 'BIRTH' but separated by 0 or more characters *or* the term 'MID' followed on the same line by 'WIFE' separated by 0 or more characters. The '|' means *or*; because of this *or* ability, it was possible to search simultaneously for up to ten phrases.

Other search phrases included:

{REPEAT AMINO}|{REPEAT BIOTIN}|{REPEAT NEONATAL}|{DATE OF COLLECTION NOT SPEC}|{SUGGEST*REPEAT}|{QUERY*?*COLLECTION} |{SEND*REPEAT}|{24 HOURS OF AGE}|{SUGGEST*CLINICAL}|{FOLLOW*UP}

and

{NAME}|{<*NEONATAL}|{< NEONATAL}|{< +NEON}| and so on.

The '<*' means anything after the beginning of a line but ending in NEONATAL and the '< N..' means NEONATAL one space after the beginning of a line and '< +N..' means NEONATAL any number of spaces after the beginning of a line. This last search was necessary because the data abstracting program would only read reports that were identified initially as being a "NEONATAL SCREEN"; but would not identify a report that was labelled as a "NEONATAL SCREE", or "NEONATAL SCRE" or variations thereof. This was necessary because the laboratory files would occasionally have an unnecessary carriage return in the middle of a word; for example

"NEONATAL SCR
EEN"

and the abstracting program did not detect these.

A final scan of all the files was done searching for the words "PEDIATRIC METABOLIC" and "PHENYLALAN". This was done because some repeats may not have been identified as a Neonatal Screen, but rather would have been part of a Pediatric

Metabolic panel, usually an amino acid screen. This scan picked up several repeat measures which would otherwise have been missed.

When relevant additional information was found, such as a request for a repeat screen, it was written in the screen file as a comment, at the end of the line describing the infant. Where the request was for a repeat specimen the comment, "SUGGEST REPEAT" was added to the screen file. Where the message was an acknowledgement of a repeat specimen, the comment "REPEAT AA" or "REPEAT NN" or "REPEAT Biotin" was added. In the instances where a T4 had been done, the comment "T4 CHEK" was added.

These search patterns exhausted the information obtained from the master file. They also required a close look at much of the lab file and allowed for a high degree of certainty that a result had not been missed. There were ten lab files for 1992, each file taking at least twelve hours of examination.

Transferring Data to a Database and More Data Cleaning

The ten completed files were then joined in one large file and entered into a REFLEX (1989) database. The variables in the database consisted of all data abstracted from the record as listed in Table 2-1. The Comment field was used extensively in the cleaning of the screen file and in cleaning and verifying the database file. All changes made to a child were recorded in this field, therefore, reversing a change was possible.

Two other variables were added at this stage. These were called CBD (Correct Birth Date), and REPEAT. CBD was used to show that the birth date had been verified against the birth record and that this name and date were correct. It later became obvious that this was not a helpful variable and it was dropped and not used in any analytic procedures. REPEAT indicated if this was a child who had a repeat blood taken and the sequence number of that repeat blood; i.e., was this the first, or second, or third, sample. A child with no repeat reports would have a REPEAT of '0'; whereas a child who had repeat reports would have an initial REPEAT of '1'. This was an iterative process and these values changed as repeat specimens were found, but in the end it permitted a quick identification of those infants who had more than one screening report.

Once in the data base, data editing continued with the key steps being:

1. Eliminating infants who had blood samples sent from out of the province, mainly the Northwest Territories and the Yukon.
2. Estimating birth dates and verifying that infants whose calculated birth date was out of scope, i.e., in 1991 or 1993, were in fact not born in 1992.
3. Detecting duplicate records and repeat tests of the same child.
4. Detecting children with identical last names and birth dates. These were commonly twins or triplets, but other children were included as well.
5. Verifying double or hyphenated names or names with apostrophes or spaces, and correcting as necessary.

Inevitably these editing stages required comparing the screen data with the 1992 birth file. Whenever a child was verified regarding name and birth date, a notation was made in the CBD field in the database to that effect. Whenever a change was made to the database, the details of the change were recorded in the "COMMENT" field in the record.

Eliminating Infants Born Out of the Province

The screening program at the WMC also serves the Yukon Territories, part of the Northwest Territories, and some border communities of Saskatchewan and British Columbia. Data for infants whose blood samples were labelled as coming from these communities were removed.

Estimating Birth date and Eliminating Out-of-Scope Birth dates

The birth date of the child was not included in the laboratory report and had to be estimated. This was done by subtracting the age of the child as recorded on the lab report from the date when the report was written. After 30 days, report age was given in months. In these instances, the age was converted into days by multiplying age in months * 45, because a child of one month could be from 30 to 60 days old. (In retrospect, in these cases it would have been better to estimate age in days as age in months*30+15.) If the child's age or the report date was unavailable, or when the calculation resulted in a negative age, the birth date was estimated as two days before the date when blood was obtained.

For infants whose calculated birth date was in late 1991 or early 1993, the name of the

child was compared with the last name of children and the last name of the mother in the birth file for birth in January or February or November or December, 1992, and included or excluded as appropriate.

Detecting Duplicate Records and Repeat Screening

The next step was to sort the file by last names and birth date and review it to detect repeat or duplicate records on the same child. This was done by searching the data file for the words "REPEAT" for "suggest repeat" and "repeat AA" or "repeat biotin", etc., "BEFORE", for "before 24h" in the COMMENT field or for "CHEK", for "T4 CHEK", in the TSH INTERP field. When a repeat record was found, the name and birth date were verified and changed as necessary and the change recorded in the COMMENT field. The name was changed if, for example, the infant had been recorded using the mother's name and this had been noted on the laboratory report. A decision was made as to whether this "extra" report was the original screen, a repeat screen or a T4 CHEK, or a duplicate report. The infant's Screen ID number (SCRID) assigned during data abstraction and linked to the original screen was then assigned to all the common records. The REPEAT field was changed from a '0', the default value, to a '1' if this were the first screen of a series, or the number of the repeat (assigned values '2' to '4' sequentially), a T4 CHEK (assigned values '5' to '8' sequentially), or a duplicate (assigned a value '9'). This ensured that an infant would not be recorded twice and that all values for the same infant would be linked by a common ID number.

Detecting children with identical last names and birth dates.

Because the match to the birth data file was keyed principally on last names and birth date, it was desirable to detect those infants who had identical last names and Birth dates and to verify as many unique characteristics as possible, particularly their given names before the match. This was seen with multiple births but it also occurred when the name was common, such as SMITH. In the original data preparation, when multiple births were found, the fact was recorded in the FirstName field as "TW A" or "TW B", or "TW C", in addition to any proper name that was there. While searching for repeat records, a search also was made in the FirstName field of the database for "TW" and the last names, first

names and birth dates verified whenever possible.

To help identify incidental identical names and Birth dates, a database report was created with LastName, FirstName, ID number, Blood Source, and Birth date. This report was then printed to a computer file and sorted by last names and birth date. This sorted file was then used as a source to find the duplicates in the database file and compare and verify names and Birth dates with the birth file like that described above.

The names and birth date of children with common surnames, e.g., Smith, Wong, and Taylor were also verified. By doing this it was possible to detect and correct many infants whose screening form had used the mother's last name. Overall, 12558 infants had their names verified.

Verifying double or hyphenated names

Some infants had names with apostrophes or spaces in them, and some infants had double names, either joined together or separated by spaces. To sort the data before matching, eliminating spaces, hyphens, and apostrophes in last names and in community names, was necessary so that the sort would list cases in both the screen and birth files in the same order. In like manner, it was necessary to verify that two last names were a hyphenated name and not the surnames of the parents, only one of which was recorded in the last name field of the birth file, or, if both were recorded, that they were recorded in the same sequence. Thus SPADY DAWSON could refer to a child whose mother was DAWSON and father was SPADY, or vice versa, or a child whose name was DAWSON-SPADY or SPADY-DAWSON. Thus, all last names with spaces, hyphens, or apostrophes were searched for and verified against the birth file. In instances where no match could be made, no change was made to the database, otherwise spaces were removed and names joined.

Obtaining and Preparing 1992 Birth Data

The data obtained from Alberta Registries were transferred to a data base and prepared for use in the matching procedure. Preparation consisted mainly of eliminating spaces from names and communities and eliminating 17 duplicate records. These had been detected by searching for children with the same last name, first name, date of birth,

mother's last name, address, and hospital of birth. As well, each birth record was assigned a unique ID number (VID). The edited birth data file was subsequently used to create an SPSS file in preparation for the matching procedure. Birth weight data were recoded to exclude those infants with a birth weight of less than 400 g, a weight which is incompatible with life and in which there was a reasonable probability of the data having been entered incorrectly.

APPENDIX 2: THE MATCHING PROCEDURE

Preparing for the Match

The raw screen data were sorted and those relevant to 1992 were imported into a Paradox® database file. In like manner a Paradox® file was created for the 1992 birth data as obtained from Alberta Registries. There were 42686 individual screening records and 42392 birth records available to match against.

Eliminating Duplicates

Before beginning the match, the screen data file was stripped of duplicate reports and repeat screenings. This was a multi-step process. The common feature in each step was to make the variable(s) of interest a key index. Where duplicate values for a variable existed, e.g. duplicate ID numbers, the data base program would remove them from the main data file and store them in a separate 'key-violation' file.

The first strip removed cases with identical screen ID numbers. The decision as to which of the two (or more) records should be in the master file was determined by the date when blood was taken, the earlier date being considered the original record. When both records had the same blood date and time, which happened commonly when a T4 was measured on the same sample, the T4 results was added to the primary record in a field "TIGHTEN". In all other events when the same blood date and time occurred the secondary record was the one which would have been a secondary investigation.

The second strip detected children with identical last names and hospital patient ID numbers. The third strip keyed on the child's last name, first name, gender, estimated date of birth, and hospital site number. The hospital site number was assigned by Alberta Health to uniquely identify hospitals in the province. During data abstraction, extra numbers were added so as to identify laboratories within the various cities that were not otherwise uniquely identified.

In a fourth strip the key variables were gender, LastName, first name, estimated birthdate, and community, where community reflected the city where the blood was taken. This strip reflected the fact that a child might have had a second blood taken at an outside lab and the hospital site number would then be different. The variable

"community" was created by linking the hospital numbers and their associated blood sources. The data file was then edited to add the name of the community of birth to each record. In the same manner this variable was added to the 1992 vital statistics birth data file. The birth data were queried to determine the frequencies of each hospital number and the associated communities.

The first strip yielded 1842 cases, the second strip 22 cases, the third strip 58 cases, and the fourth strip only 14 cases. The reason that the pickup after the first strip was so modest is likely because most of the repeat visits had been detected in the initial data cleaning.

The screen data file was then queried to establish how many children had the same last name, estimated birth date, and community. There were 1227 instances. To determine if any of these were duplicates, the screen and birth data files were linked and those cases were examined where two or more children with the same last name and who were born on the same date in the same community and a judgment made as to if they were, or were not, the same child. Children deemed identical were assigned the same ID and their names, usually their first names, were changed to be the same. When this was finished, a fifth strip, using gender, last name, first name, estimated date of birth, and community was done. This picked up 48 more cases.

Duplicates were also searched for in the screen data by reviewing individually all the cases of children whose birth was estimated to be in the same week, in the same community, and with the same last name. These were children where there was the possibility that a mistake had been made in estimating their birth date. It quickly became obvious that this exercise might find a few names but it was a very inefficient use of time; therefore, a sixth strip, using Birth Week as the new variable, was done and picked up five more cases.

The total number of duplicate cases detected using these stripping procedures was 1989 cases. There were a few more duplicates detected as the match progressed, but the vast majority were found in these initial strip procedures.

The Match Process

With the completion of the stripping procedures, the match on the remaining data was started. To avoid redundancy, please refer to Table 2-3 which summarizes the final accounting of the match. It lists the criteria used for each of the matching steps, the confidence level assigned, and the resulting yield. It is a summary document and does not reflect the many interim steps taken, and the revisions to individual matches made during the match process.

The first set of matches consisted of a series of MATCH procedures in SPSS. After four runs, 37377 records were matched. These matches were highly deterministic and depended on the accurate listing of the infants' gender, the infant's or the mother's last name, first name, date of birth, and birth place.

The First Hand Matches

The above matches required the child's last name in the screen file to be identical to either the child's or the mother's surname in the birth file. To find cases with potential spelling errors and thus with names similar but not identical to those in the birth data, the screen data had to be searched by hand for similar but not identical spelling. To do this, the remaining screen data were linked to the remaining birth data and the birth data were searched for names that were 'like' a name in the screen file but which were misspelled; e.g. LEFEBuRE instead of LEFEBvRE. These matched cases were removed from the active files and placed in a separate file and eventually added to the master file. Two duplicates were also found.

The Second SPSS Matches

In doing the above procedure it was apparent that many links did not occur in the SPSS MATCH because of simple problems such as the wrong gender being recorded. The hand procedure was also labour intensive and it became apparent that matching could proceed effectively using only last name, date of birth, and community, but NOT gender. Therefore a second series of SPSS matches was done and a further 1064 records were matched.

The Second Set of Hand Matches

The remaining records transferred to a database file and an attempt made to hand match as many as possible. The method used was to sort the screen and birth data files by date of birth and community and then search within the community for appropriate names, by visual inspection and by searching for 'like' names. This worked fairly well and there was a significant amount of matching especially in the smaller centres.

The Match on Previously Matched Data

Upon completion of the second set of hand matches, the matched screen data file was matched to a file containing the remaining unmatched screen data. This 'match on previously matched data' was designed to detect cases where repeats had been done, but where no record had been made that this particular screen was a repeat. A further 343 cases were found where the most likely match was with a child already matched.

The Final Match Process

At the end of these series of matches, 823 unmatched screen records remained. These records were matched using the most crude of links. An attempt was made to match by using Soundex codes for last names, plus the birth date or week and the community. This resulted in 24 more matches. Most of the remaining cases were matched by matching a child's birth week to a previously unmatched birth record from the same week and the same community. The remaining children were matched by hand linking a screened child to a child whose birth record had the closest birthdate and who was born in a nearby community, or was born in a major referral centre.

This ended the matching procedure; each screen record was linked with a birth data record or another, matched, screen record. The next steps were: (i) to assign a level of confidence to each match; (ii) to ensure that any duplicate cases found had the same ID number and had a sequence assigned to their measurements; and, (iii) to review the data for inconsistencies, such as having blood taken before birth, that reflected potential errors in matching or in data entry.

Assigning Levels of Confidence

Each matched record was assigned a level of confidence describing an estimate of the

accuracy of the match. Table 2-3 also summarizes the final levels of confidence assigned. Initially there were grades of confidence from one to nine but on reflection it was decided to state that a match was one of good confidence and thus was a very good match (GM) or was an uncertain match (UM). In general, those records matched in Series 1, 2, and 4, and some of Series 3 (Table 2-3), were considered good matches. The remaining matched records were considered uncertain matches.

Reconciling Duplicate Cases

Each match step described above resulted in an output file. These output files were united to make one master output file containing the VID, SCRID, confidence level of the match, and the file source. It did not include known repeat screens or duplicate cases. To continue the analysis and match, the screen records that had been stripped before the match were added to the screen data and the master output file was linked to this to make a complete screen file. The resultant file then consisted of a link to the birth data (VID), a link to the screen file (SCRID), the source of the match, and the confidence of the match, plus all the screen data. This was a necessary step before determining the sequencing of duplicate screens.

The first step was to reconcile the duplicate cases that had been stripped from the screen file before the match started, plus the few extra that had been found incidentally during the match procedure as well as the duplicate cases that had been found during the match on matched data (i.e. when unmatched screen data were matched with previously matched screen data). The intent here was first to ensure that a duplicate case had the same screen ID (SCRID) and secondly to prepare the data so that the sequence of duplicate screens could be determined. A unique counter number was then assigned to each case. This was done because a unique identifier was required to have as a key variable for the data base and the SCRID would not be satisfactory, because the duplicate cases obviously would have the same SCRID. Then the data were sorted and the SCRID's were reconciled by hand so that each child had only one SCRID, even though the child might have more than one record. This SCRID thus became the child's unique identifier, as opposed to the record's unique identifier.

The records that were a repeat sample or were duplicates were then extracted from the main screen file. This file of duplicates was linked up with the matched screen data file and the sequence of measurements was determined and the record edited appropriately. This involved examining the two (or more) records to determine the time and date of the blood taking and assigning a value of '1' to the variable REPEAT for the first measurement and a value of '2', or whatever, to the variable REPEAT for the second or later measures. If the repeat measurement was a T4 CHEK, a value of '5' was assigned to the first measurement and a value of '6' or '7' to the second or third measurement. Where the duplicate was in every respect a duplicate report only, and not a repeat measure, the value of '9' was assigned. In every instance, if a child had only one measurement taken, the REPEAT variable was assigned a value of '0'. In this manner, by searching the database for a REPEAT value of '1' it was possible to quickly find all cases where there was a repeat screen. A subsequent iteration of the file removed all T4 checks that had been obtained from an original sample. These results were placed in a field on the main screen data file. This was done because a T4 obtained from an original sample was not a repeat, and the possibility of analyzing it as such had to be minimized. Only genuine repeat T4 checks were then assigned a value of '5' in the T4 check field.

Looking for Inconsistencies

Two problems soon became obvious. The first was the discovery that a lot of children were born who had no record of being screened. The second was that sometimes children were recorded as having had blood taken before they were born.

Missing Children

The first problem was detected after an early attempt to determine completeness of coverage. There were 42686 screen data records describing the screen results of 40359 children. But 42392 children were born, an excess over screening of 2033 children. Of these, a graphical analysis of the data showing match status plotted against date of birth showed that an estimated 655 children were missing from a time period spanning days 159 to 172 of 1992 and an estimated 548 children were missing from a period spanning days 289 to 300. The estimate was determined by averaging the daily number of infants not

accounted for during the rest of the year and multiplying by the number of days missed and subtracting the "excess" missed infants from these figures. The laboratory information manager was informed. It transpired that there were some laboratory problems during these periods. A backup laboratory file was searched and the relevant data obtained. These new data were abstracted, cleaned, edited, and matched in a manner similar to that described above for the main data set and 1194 more children were added to the master screen data file.

Aberrant Birth and Blood Dates

A second problem was that of verifying the date of birth and the date when blood was taken. The age at blood taking was calculated by linking the birth data files and the screen data files and calculating the number of days between date of birth, obtained from the birth file, and date of blood, obtained from the screen file. After doing this and sorting the output it was obvious that some children supposedly had blood taken days to months before they were born, and yet the confidence level of the match was "good". Each of these records was examined individually and a decision made as to how to deal with the problem.

In 79 instances the record, but not the match, was changed. These usually were errors in data entry. This opinion was arrived at by ranking the files in terms of lab number and then examining those cases with negative ages at blood taking. The sequence of lab numbers drew attention to the fact that the date entered was an inversion of numbers; e.g. 12 instead of 21, or the date was out by a month, or out by a year. Rarely, an arbitrary decision was made. If so, the reasons were a good link on last name and first name and place and surrounding date values in the raw data file (when sorted by lab number). Only rarely was the match changed. For example, subject "A" was randomly matched and subject "B" was a "good match" child with the same name/sex born on the same date. "A" had a before 24h comment and "B" had a blood age of 20 days. On review it was obvious that these were the same child. The VID that had been attributed to "A" was then removed and the SCRID changed to match that of "B".

Regional Assignment

The recently created Health Authority Regions were incorporated into the database to allow examination of potential regional problems. The boundaries were those set as of summer 1995 and the regions were linked to hospitals and communities using data provided by Alberta Health and by reference to a map of Alberta. For homes out of province the region was coded as 21.

Dummy Records

At times, during data abstraction, records with Repeat AA comments were found that had no earlier record linked to them. In this instance, the screen file was examined to see if there was another record with the same name and SCRID. If so the file was amended appropriately. The birth file was then examined to determine the mother's name for the child and to see if there was a record made in her name rather than the child's surname. If so, the necessary corrections were made. In 63 cases where the above efforts did not work, a REPEAT value of 2 was assigned and the necessary match made with the birth file. At the same time, a dummy initial record was created, in order to ensure proper calculation of the estimate of coverage and to obtain demographic data regarding these infants. Of the 63 'dummy' infants, 49 were in the good match group.